

UNIVERSIDAD AUTÓNOMA DE MANIZALES

MAESTRÍA EN GESTIÓN Y DESARROLLO DE PROYECTOS DE SOFTWARE



DESCUBRIMIENTO DE CONOCIMIENTO EN UNA EMPRESA DE
OUTSOURCING DE TI DE LA CIUDAD DE MEDELLÍN APLICANDO TÉCNICAS
DE MINERÍA DE DATOS QUE PERMITA IDENTIFICAR POTENCIALIDADES
EN EL ÉXITO DE LOS PROYECTOS DE DESARROLLO DE SOFTWARE

MAURICIO GALLEGO GALLEGO

MANIZALES, CALDAS – COLOMBIA

OCTUBRE 2014

DESCUBRIMIENTO DE CONOCIMIENTO EN EMPRESA DE *OUTSOURCING* DE
TI DE LA CIUDAD DE MEDELLÍN APLICANDO TÉCNICAS DE MINERÍA DE
DATOS QUE PERMITA IDENTIFICAR POTENCIALIDADES EN EL ÉXITO DE
LOS PROYECTOS DE DESARROLLO DE SOFTWARE

MAURICIO GALLEGO GALLEGO

Informe final

Msc. Javier Hernández Cáceres
Director: Temático y Metodológico
Docente Investigador

UNIVERSIDAD AUTÓNOMA DE MANIZALES
MAESTRÍA EN GESTIÓN Y DESARROLLO DE PROYECTOS DE SOFTWARE
MANIZALES

2014

Tabla de contenido

Resumen	9
1. INTRODUCCIÓN	10
2. REFERENTE CONTEXTUAL	14
2.1. Descripción del Área Problemática.....	14
2.2. Antecedentes	17
2.2.1 Antecedentes técnicos y temáticos.....	17
2.2.2 Antecedentes temáticos en Colombia	25
2.2.3 Antecedentes metodológicos en Colombia.....	28
2.3. Justificación.....	32
2.4. Formulación del problema.....	34
2.5. Objetivo	35
2.5.1. Objetivo General	35
2.5.2. Objetivos Específicos.....	35
2.6. Resultados esperados	36
3. ESTRATEGIA METODOLÓGICA	37
3.1. Metodología	37
3.1.1. Comprensión del negocio.....	39
3.1.2. Comprensión de datos.	40
3.1.3. Preparación de datos.	40
3.1.4. Modelado.....	41
3.1.5. Evaluación.	42
3.1.6. Despliegue.....	42
3.2. Pruebas	43
3.3. Presupuesto	43

4.	DESARROLLO	44
4.1.	Referente teórico	44
4.1.1.	Dato, información y conocimiento.....	44
4.1.2.	Minería de datos.....	45
4.1.3.	Técnicas de minería de datos.....	46
4.1.4.	Descubrimiento de conocimiento en bases de datos.....	51
4.2.	Factores que influyen en la inyección de defectos de desarrollo de software. .	53
4.3.	Comprensión del negocio	68
4.3.1.	Objetivos institucionales.....	68
4.3.2.	Evaluación de la situación.....	68
4.3.3.	Objetivos de la minera de datos.....	71
4.4.	Comprensión de los datos.....	71
4.4.1.	Recolección inicial de datos.....	71
4.4.2.	Descripción de datos.....	77
4.4.3.	Exploración de datos.....	86
4.4.4.	Verificar la calidad de los datos.....	98
4.5.	Preparación de los datos	99
4.5.1.	Selección de datos y limpieza de datos.....	99
4.5.2.	Construcción de datos.....	99
4.5.3.	Integración de datos.....	100
4.6.	Modelado	100
4.6.1.	Seleccionar la técnica de modelado.....	100
4.6.2.	Generar el diseño de las pruebas.....	101
4.6.3.	Construcción del modelo.....	102
4.6.4.	Evaluación del modelo.....	105

4.7.	Evaluación.....	110
4.7.1.	Capacidad de los desarrolladores.	110
4.7.2.	Conocimiento del dominio.	115
4.7.3.	Comunicación.	121
4.7.4.	Composición del equipo.	125
4.7.5.	Complejidad del producto.	132
4.7.6.	Madurez de gestión de proyectos.	136
4.7.7.	Selección de factores.	141
4.8.	Despliegue.....	150
4.9.	Análisis de resultados	151
5.	CONCLUSIONES	153
6.	RECOMENDACIONES	158
	BIBLIOGRAFÍA.....	159
	ANEXOS	165

Lista de figuras

Figura 1 Factores críticos de éxito en proyectos de integración de datos.	19
Figura 2 Matriz de adyacencia.	20
Figura 3 Centralidad de los Factores.	20
Figura 4 Tabla comparativa algoritmos de minería de datos	23
Figura 5 Metodología CRISP-DM.....	39
Figura 6 Fases del proceso KDD	52

Lista de tablas

Tabla 1. Operacionalización de variables	78
Tabla 2. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría capacidad de los desarrolladores.....	87
Tabla 3. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría conocimiento del dominio.	89
Tabla 4. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría comunicación.....	90
Tabla 5. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría composición del equipo.....	92
Tabla 6. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría complejidad del producto.	94
Tabla 7. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría madurez de la gestión de proyectos.....	95
Tabla 8. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría colaboración del equipo.....	97
Tabla 9. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría ejecución del proyecto.....	98
Tabla 10 Variables en la ecuación del modelo	104
Tabla 11 Resumen del modelo de la regresión logística bivariado	106
Tabla 12 Variables que no están en la ecuación.....	107
Tabla 13 Pruebas ómnibus sobre los coeficientes del modelo	108
Tabla 14 Prueba de Hosmer y Lemeshow	108
Tabla 15 Matriz de correlaciones entre los factores que componen la categoría capacidad de los desarrolladores.....	111
Tabla 16 Cuadro de comunalidades capacidad de los desarrolladores	113
Tabla 17 Varianza total explicada capacidad de los desarrolladores	114
Tabla 18 Matriz de correlaciones entre los factores que componen la categoría conocimiento del dominio	116
Tabla 19 Cuadro de comunalidades conocimiento del dominio	118

Tabla 20 Varianza total explicada conocimiento del dominio	119
Tabla 21 Matriz de correlaciones entre los factores que componen la categoría comunicación.....	122
Tabla 22 Cuadro de comunalidades comunicación	123
Tabla 23 Varianza total explicada comunicación.....	124
Tabla 24 Matriz de correlaciones entre los factores que componen la categoría composición del equipo.....	126
Tabla 25 Cuadro de comunalidades composición del equipo	128
Tabla 26 Varianza total explicada composición del equipo.....	129
Tabla 27 Matriz de correlaciones entre los factores que componen la categoría complejidad del producto.	133
Tabla 28 Cuadro de comunalidades complejidad del producto	135
Tabla 29 Varianza total explicada complejidad del producto	135
Tabla 30 Matriz de correlaciones entre los factores que componen la categoría Madurez de gestión de proyectos	137
Tabla 31 Cuadro de comunalidades madurez de gestión de proyectos.....	139
Tabla 32 Varianza total explicada madurez de gestión de proyectos.....	140
Tabla 33 Estadísticos de colinealidad.....	143
Tabla 34 Matriz de correlaciones entre los factores seleccionados para la construcción de un modelo multivariado.....	144
Tabla 35 Cuadro de comunalidades factores análisis multivariado	146
Tabla 36 Varianza total explicada factores análisis multivariado.....	147
Tabla 37 Tabla de clasificación	153

Resumen

Lograr el éxito en los proyectos de desarrollo de software es una tarea a la que día a día se están enfrentando las empresas productoras de software, estudios realizados por diferentes organizaciones demuestran que la tasa de éxito de los proyectos esta alrededor del 32%, y de proyectos fallidos alrededor del 24% (Eveleens & Verhoef, 2010), siendo la diferencia de estos una cifra relativamente baja lo que conlleva a que se deba evaluar el tema y se piense en tomar medidas; el éxito de los proyectos depende de la sincronización que se pueda tener entre diferentes factores como los humanos, los técnicos, los económicos, tiempo, alcance, entre otros.

En este sentido, surgió la necesidad de poder generar conocimiento útil a partir del análisis de datos históricos de proyectos exitosos y fallidos de una empresa de *outsourcing* de TI aplicando técnicas de minería de datos, con el fin de obtener patrones que permitieran a la organización tomar decisiones y guiar los proyectos de software a la consecución del éxito.

La investigación fue desarrollada siguiendo la metodología CRISP-DM, que es una metodología utilizada como referente para los proyectos de minería de datos, para la aplicación de las técnicas se utilizó la herramienta de software SPSS en su versión trial.

1. INTRODUCCIÓN

Con el paso de los años y la evolución tecnológica, la innovación y la facilidad de agilizar los procesos operativos diarios del ser humano; el mercado mundial ha convergido en gran porcentaje al desarrollo de proyectos de software (UNCTAD, 2012). Gran cantidad de empresas internacionalmente han extendido su nombre gracias al éxito de sus sistemas informáticos que soportan el *core* de importantes compañías mundiales. Pero a pesar de que se encuentran en el tope del éxito, no todas sus iniciativas de software han alcanzado y obtenido buenos resultados; de allí que incluso en esta era donde abundan los marcos de trabajo, estudios, procesos y metodologías entre otros para garantizar la consecución plena de sus objetivos permanece la duda de cómo garantizar el éxito de un proyecto.

En el presente trabajo, se abordará desde la gestión del conocimiento (Mas Machuca & Martínez Costa, 2008) técnicas para la obtención de información y datos dentro de la empresa consultora, que permitan determinar factores de éxito específicos para los diferentes proyectos de software, bien sea por la cantidad de personal requerido, las habilidades necesarias para afrontar los proyectos, la experiencia en las herramientas requeridas para el desarrollo de los requisitos, cualidades de los trabajadores entre otros.

Por medio de la minería de datos se busca conocer las razones y consecuencias de los proyectos principales de la compañía que hayan sido críticos en su progreso y crecimiento; pero sin dejar atrás aquellos que hayan sido un punto de inflexión donde sus metas hayan sido obstruidas.

A partir de la información obtenida, se pudo entonces identificar patrones que permitirán ubicar las iniciativas de desarrollo en diferentes ambientes con el fin de procurar buenos resultados de los mismos, así mismo, se realizó un análisis según el resultado de los modelos obtenidos en el estudio tomando como referencia la historia y experiencias de los casos de éxito y fracaso de la empresa.

Esta investigación se realizó según los lineamientos expuestos por la metodología CRISP-DM, ya que es utilizado como referente para el desarrollo de proyectos de minería de datos a nivel mundial.

En el desarrollo del presente trabajo se presentaron diferentes situaciones que implicaron realizar algunos cambios y ajustes a la propuesta inicialmente presentada.

El título original de la investigación hacía referencia a: Descubrimiento de conocimiento en las bases de datos de una empresa de *outsourcing* de TI de la ciudad de Medellín aplicando técnicas de minería de datos que permita identificar potencialidades en el éxito de los proyectos de desarrollo de software; cuando se presentó la propuesta del proyecto se contó con la participación del líder de arquitectura de la empresa en ese entonces, con el cual se hizo una revisión preliminar de la información que se debía recolectar de los proyectos y quien en su momento nos indicó que se tenía toda la información de los proyectos almacenadas en bases de datos; esta persona posteriormente renunció a la compañía, y una vez se realizó la reunión con la persona que realizó el reemplazo se detectó que la base de datos que debían contener los datos históricos de los proyectos no contenía la información necesaria para hacer el respectivo estudio, las personas encargadas de suministrar los datos de los proyectos no lo estaban haciendo, es decir, no

estaban cumpliendo con el proceso de la compañía, por tal motivo se optó por recopilar los datos de los proyectos históricos a partir de encuestas, lo que implicó reformular el título de la investigación por: Descubrimiento de conocimiento en una empresa de *outsourcing* de TI de la ciudad de Medellín aplicando técnicas de minería de datos que permita identificar potencialidades en el éxito de los proyectos de desarrollo de software.

Una vez se desarrolla el primer objetivo de la investigación: Realizar una búsqueda bibliográfica sobre los factores que influyen en la inyección de defectos de desarrollo de software; y al realizar el análisis y selección de los factores que se iban a estudiar, se detecta que dichos factores ya contaban con una clasificación o segmentación, la cual fue realizada por los autores consultados, adicionalmente ya se contaban con dos grupos definidos éxito y fracaso (para hacer referencia a factores de proyectos exitosos y factores de proyectos fracasados) esto conllevó a realizar un cambio en los objetivos iniciales y resultados esperados, por los objetivos presentados en este informe que de igual forma suministraron valor a la investigación; de esta forma se suprimió el objetivo: “Segmentar el número de factores por medio de un análisis de *clúster*, y se adicionaron los objetivos”:

- Realizar una descripción e interpretación de los datos de los proyectos de desarrollo de software estudiados.
- Construir un modelo multivariado que permita predecir el éxito o fracaso de un proyecto de desarrollo de software en la empresa de *outsourcing* de TI de la ciudad de Medellín.
- Realizar una simulación del modelo multivariado con los datos de los proyectos de desarrollo de software objeto de estudio.

De igual forma se siguió la recomendación realizada por el ente evaluador del proyecto en cuanto al objetivo “Reducir el número de defectos residuales en proyectos de desarrollo de software, utilizando extracción de conocimiento de la base de datos de la empresa de *outsourcing* de TI de la ciudad de Medellín”, de darle un manejo diferente por las implicaciones de alcance y validaciones que este podría requerir, por tal motivo, el manejo de este objetivo propuesto en la fase del proyecto se presentó en las conclusiones.

2. REFERENTE CONTEXTUAL

2.1. Descripción del Área Problemática

El estudio del éxito de los proyectos de software ha sido un tema de gran atención desde hace más de una década, en la que se ha tratado de analizar los factores que colaboran al cumplimiento completo de un desarrollo y la manera como las empresas tratan de evitar posibles fracasos. Un ejemplo fue el análisis realizado por Laurenz Eveleens & Verhoef (2010) sobre las estadísticas del Standish Group de proyectos exitosos donde evidenciaba que el crecimiento de efectividad de los proyectos entre 2000 y 2004 había sido solo de un 1%, situación que mejora al 2009 con un 4%. Cabe aclarar que con el paso de los años los factores tenidos en cuenta para el estudio han estado ampliándose para hallar con más efectividad una respuesta (Carroll, 2013). De 1999 donde se consideraban solo 5 elementos, para el 2012 han sido duplicados.

Un análisis desde otro flanco, incluyendo la cara oscura en el desarrollo de proyectos de software, la variación respecto a fracasos o rescate de los mismos si registra grandes diferencias. Como referencia entre los años 1994 a 2012, los proyectos fallidos pasaron de un 31% a 18%, y proyectos exitoso registraron una diferencia de 23% de más respecto a 1994 donde equivalían al 16% (Nizam Nasir & Sahibuddin, 2011), (Standish Group, 2013).

Los resultados son alentadores, sin embargo no van en tendencia creciente, registrando años en donde el éxito, fracaso o mediano cumplimiento aumenta o disminuye variablemente.

Si bien, aunque cada entidad consultora ha ido perfeccionando la respuesta a los factores de consecución de los proyectos de software, todas convergen en tres principios principales: tiempo, costo y alcance (Linberg, 1999).

Es claramente entendido que un proyecto de software es exitoso cuando el presupuesto estimado no es sobrepasado, el tiempo de entrega propuesto se cumple con exactitud y han cumplido con los objetivos propuestos; sin embargo es necesario dar atención a demás elementos que hacen parte para que estos componentes se alcancen.

Un proyecto de software está conformado básicamente por desarrolladores, analistas de pruebas, analistas de requerimientos y jefes de proyecto, de quienes se espera un buen entendimiento, trabajo en equipo, disponibilidad, liderazgo, motivación, experiencia y *skills* aptos para afrontar los retos. Podría considerarse entonces que si una de estas cualidades o características está ausente o falla dentro del grupo de trabajo, alguno de los tres principios de éxito perderá los lineamientos. Fallas como: pobre estructura organizacional, falta de liderazgo, falta de apoyo del nivel gerencial, falta de esfuerzo, choques de personalidades; uso inefectivo de métodos de desarrollo de software, procesos de negocios y asignación de recursos inapropiados, gestión de proyectos y herramientas de seguimiento inadecuados (Pereira R, Cerpa T & Rivas M, 2004).

Además de todas las propuestas que se encuentran en la teoría, es importante contar con herramientas que permitan obtener ideas simples de lo que fundamentalmente se requiere

para entablar buenos cimientos que colaboren a garantizar un proyecto exitoso; de aquí el poder basarse en juicios de expertos o iniciativas anteriores que hayan cumplido con las necesidades, tanto para el proveedor como el cliente.

Dentro de empresas de gran magnitud, cuya fuente de sostenibilidad es el desarrollo de sistemas de información para clientes de altas exigencias, aprovechar el conocimiento de experiencias pasadas es un punto a favor para procurar no repetir errores y mitigar cada vez más un posible fracaso (CHATZTEL, 2003) citado por (Mas Machuca & Martínez Costa, 2008).

2.2. Antecedentes

2.2.1 Antecedentes técnicos y temáticos.

2.2.1.1 *Factores críticos de éxito de la industria del software y su relación con la orientación estratégica de negocio: un estudio empírico-exploratorio.*

González & Rodenes (2007) realizaron un estudio empírico exploratorio aplicado a la industria del software en el país de México, donde se plantearon buscar los factores críticos de éxito de la industria del software y su relación con la orientación estratégica del negocio (coste y diferenciación); para tal investigación y para efectos de obtener una clasificación de las empresas participantes en el estudio, eligieron la técnica estadística de análisis de conglomerados (*clúster analysis*) y como técnica complementaria para dar validez a la clasificación el análisis discriminante; así mismo, para contrastar el grado de relación de los FCE con la orientación estratégica de negocio utilizaron el análisis de correlación bivariada y análisis de regresión.

Como resultado de la aplicación de las técnicas de análisis de conglomerados obtuvieron dos conglomerados: uno con 30 empresas para el grupo de estrategia por costes (44%) y con 38 (56%) empresas para el grupo de estrategia por diferenciación.

Como resultado de la aplicación de la técnica de análisis discriminante se obtuvieron dos conglomerados: uno con 30 empresas para el grupo de estrategia por costes y otro con 38 empresas para el grupo de estrategia por diferenciación, corroborando así el resultado

obtenido con la aplicación de la técnica de análisis de conglomerados, tanto en número de conglomerados, casos y variable de clasificación (porcentaje de productos o servicios especializados ($p < .001$))

Para analizar la relación entre los grupos de orientación estratégica y los factores críticos de éxito de la Industria del software, se aplicó la técnica de análisis bivariado que, según Rowland (2004) citado por González & Rodenes (2007), es una prueba que puede ser usada para analizar la diferencia entre grupos o bien para detectar una asociación, correlación o relación entre dos variables.

De los cinco factores críticos de éxitos identificados (Apoyo del Gobierno, Capital Humano, Calidad, Marketing e Innovación), los únicos que mostraron tener una relación de manera positiva y significativa con el éxito de la orientación estratégica de negocio de acuerdo al análisis realizado fueron Capital Humano, Calidad e Innovación.

2.2.1.2 Modelado y análisis de los factores críticos de éxito de los proyectos de software mediante mapas cognitivos difusos.

Leyva, Rosado & Febles (2010) indicaron que la determinación y análisis de los factores críticos de éxito en los proyectos de software contribuye a que las organizaciones dedicadas al desarrollo de software centren su atención en los factores fundamentales para ser exitosas. Por tal motivo realizaron un trabajo el cual tenía como objetivo proponer una metodología basada en los mapas cognitivos difusos para la formalización y el análisis de los factores críticos de éxito. La metodología propuesta consta de actividades que permiten la identificación de los factores críticos de éxito, la

construcción de un Mapa cognitivo difuso que los represente, y finalmente el análisis del mismo.

Con el fin de sustentar la metodología, Leyva, Rosado & Febles (2010), realizaron un caso de estudio en una organización de desarrollo de software en Cuba; para la determinación de los factores críticos de éxito se basaron en la literatura existente, factores utilizados en proyectos tecnológicos afines con los almacenes de datos, y los factores especificados por entidades externas que se dedican a la temática:

Figura 1 Factores críticos de éxito en proyectos de integración de datos.

Id	Factor
N1	Apoyo de la alta gerencia
N2	Participación de los usuarios
N3	Conocimientos técnicos de los usuarios
N4	Definición del alcance y prioridades
N5	Calidad de las fuentes de datos
N6	Entrenamiento de los usuarios
N7	Definición de requisitos
N8	Tiempo
N9	Compromiso del cliente con la solución
N10	Disponibilidad tecnológica

Nota fuente: Tomado de Leyva, Rosado & Febles (2010) (p. 43).

Para el desarrollo del mapa cognitivo difuso, obtuvieron 11 mapas cognitivos difuso, que después de integrarlos obtuvieron la siguiente matriz de adyacencia:

Figura 2 Matriz de adyacencia.

Id	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
N1	0.00	0.35	0.00	0.20	0.00	-0.10	0.10	-0.45	0.35	0.20
N2	0.80	0.00	0.00	0.50	0.00	0.35	0.29	0.00	0.65	0.00
N3	0.35	0.35	0.00	0.00	0.00	0.67	0.00	0.10	0.00	0.20
N4	0.65	0.89	0.80	0.00	0.00	0.10	0.65	0.00	0.55	0.10
N5	0.10	0.00	0.44	0.28	0.00	0.00	0.00	0.00	0.00	0.20
N6	0.50	0.60	-0.65	0.20	0.00	0.00	0.10	0.00	3.50	0.00
N7	0.58	0.89	0.65	0.80	0.20	0.00	0.00	0.00	0.50	0.00
N8	-0.35	-0.76	-0.65	-0.65	-0.76	0.00	0.00	0.00	-0.70	-6.50
N9	0.73	0.67	0.00	0.50	0.20	0.50	0.50	-0.50	0.00	0.00
N10	0.73	0.00	0.00	0.35	0.00	0.00	0.00	0.00	0.00	0.00

Nota fuente: Tomado de Leyva, Rosado & Febles (2010) (p. 43)

Para la realización del análisis estático, Leyva, Rosado & Febles (2010), obtuvieron el grado de entrada, de salida y calcularon la centralidad. A partir de este valor los factores críticos de éxito los ordenaron de mayor a menor:

Figura 3 Centralidad de los Factores.

Factor	Outdegree	Indegree	Centralidad	Centralidad
Alcance y P.	3.48	3.74	7.21	0.139
Participación U	4.50	2.59	7.09	0.136
Compromiso del cliente	3.10	3.60	6.70	0.129
Apoyo AA	4.78	1.75	6.53	0.126
Tiempo	1.05	4.52	5.57	0.107
Definición de R.	1.72	3.61	5.33	0.102
Conocimientos T.U	3.19	1.67	4.86	0.093
Entrenamiento U	1.72	2.40	4.12	0.079
Disponibilidad tecnológica	1.35	1.08	2.43	0.047
Calidad Fuentes de datos	1.16	1.02	2.17	0.042

Nota fuente: Tomado de Leyva, Rosado & Febles (2010) (p. 44)

Dentro del sistema formado por los factores críticos de éxito, los tres factores más importantes en orden descendente son: la definición del alcance y prioridades del proyecto, la participación de usuario final, y el grado de compromiso de los clientes.

2.2.1.3 The optimization of success probability for software projects using genetic algorithms.

Según Reyes, Cerpa, Candia-Véjar & Barden (2011) por lo general el proceso de desarrollo de software se ve afectado por muchos factores de riesgo que pueden causar el fracaso y la pérdida del control del proyecto, por lo que estos riesgos deben ser identificados y mitigados por los jefes de proyecto. Empresas de desarrollo de software están mejorando sus procesos mediante la adopción de prácticas aceptadas a nivel internacional, con el fin de evitar riesgos y demostrar la calidad de su trabajo.

El trabajo presentado por Reyes et. Al (2011) tenía como objetivo desarrollar un método para identificar los factores de riesgo más influyentes en la determinación de los resultados del proyecto. Este método también debería proponer una inversión rentable de los recursos del proyecto para mejorar la probabilidad de éxito del proyecto.

Para alcanzar estos objetivos, utilizaron la probabilidad de éxito en relación a los costos para el cálculo de la eficiencia de los resultados del proyecto. Utilizaron la eficiencia como la función física en una técnica de optimización basado en algoritmos genéticos. Logrando así, que el método maximice la salida de probabilidad de éxito de un modelo de predicción con respecto al costo. Reyes et. Al (2011).

El método de optimización lo probaron con varios modelos de predicción del riesgo de software en base a métodos bayesianos (BBN, NBC, SBC) para tipos de resultados del cliente y del líder de proyecto, que se han desarrollado sobre la base de la literatura y el

uso de los datos de una encuesta que recogía la información de proyectos internos y externos de empresas en la industria chilena de software. Estos modelos predicen la probabilidad de éxito de un proyecto basado en las actividades realizadas por el director del proyecto y el equipo de desarrollo. Según Reyes et. Al (2011) los resultados mostraron que el método propuesto es muy útil para identificar aquellas actividades que requieren una mayor asignación de recursos, y cuáles de ellas tendrán un mayor impacto en la probabilidad de éxito de los proyectos.

Results of genetic algorithm optimization (with cost restrictions).

Model	# of generations	Best fitness	Initial probability of success (%)	Ending probability of success (%)	Initial cost	End cost
BBN outcome_client	300	0.887	44.49	96.13	37	25
BBN outcome_leader	43	0.707	20.64	100	30	30
NBC outcome_client	300	0.584	1.17	84.85	216	241
NBC outcome_leader	43	0.333	2.03	48.26	411	414
SBC outcome_client	300	0.566	5.79	81.45	25	30
SBC outcome_leader	53	0.218	2.67	89.82	21	34

Results of genetic algorithm optimization (without cost restrictions).

Model	# of generations	Best fitness	Initial probability of success (%)	Ending probability of success (%)	Initial cost	End cost
BBN outcome_client	289	0.887	44.49	96.13	37	25
BBN outcome_leader	45	0.707	20.64	100	30	30
NBC outcome_client	257	0.647	1.17	97.55	216	341
NBC outcome_leader	52	0.613	2.03	98.82	411	1814
SBC outcome_client	52	0.674	5.79	97.36	25	59
SBC outcome_leader	51	0.218	2.67	89.82	21	34

2.2.1.4. Software Projects Success Factors Identification using Data Mining.

Según Yousef, Gamal, Warda & M.Mahmoud (2006) el éxito de un proyecto se define como el logro de los objetivos y características del proyecto, dentro del costo y tiempo asignado, indicaron que son muchos los factores que influyen en el éxito de un proyecto pero se centraron principalmente en la recolección de datos que intervienen en los

factores de éxito: requisitos, patrocinador del proyecto, y los clientes, de 40 proyectos de desarrollo de software de la industria egipcia.

Para el análisis de la información utilizaron las siguientes técnicas de minería de datos: redes neuronales, reglas de asociación, *clustering*, el clasificador Naive Bayes y arboles de decisión que les permitiera descubrir características y reglas que rigen el éxito y fracaso de los 40 proyectos.

Los resultados de los autores mostraron que cada algoritmo de minería de datos tiene un punto fuerte para proporcionar conocimientos y hacer predicciones sobre el éxito de los proyectos.

Los autores indicaron que muchas son las medidas utilizadas para comparar la habilidad de realizar predicciones de las técnicas de minería de datos, para el presente estudio utilizaron "Gráfico de elevación", "Matriz de clasificación" y "Precisión estimación", obteniendo como resultado la siguiente tabla comparativa:

Figura 4 Tabla comparativa algoritmos de minería de datos

MINING ALGORITHMS MEASURES COMPARISON

Algorithm	Lift Chart/ Classification Matrix	Estimation Accuracy
Association	58%,38%	39%,17%
Naïve Bayes	82%,74%	48%,51%
Neural Network	46%,51%	43%,39%
Clustering	69%,69%	53%,49%
Decision Tree	51%,61%	N/A

Nota fuente: Tomado de Yousef, Gamal, Warda & M.Mahmoud (2006) (p. 452)

Yousef, Gamal, Warda & M.Mahmoud (2006) indican que los valores de la tabla muestran que a pesar de la cantidad limitada de registros (sólo 40), la precisión de las técnicas de minería son aceptables especialmente el clasificador Naives Bayes, las redes neuronales y los clústeres. También indican que los valores de "Precisión estimación" son menores en comparación con "Gráfico de elevación" y "Matriz de clasificación" debido a la utilización de sólo 7.500 de los datos como un conjunto de prueba; de igual forma afirmaron que aunque el árbol de decisión destacó el orden de importancia de los factores de éxito o fracaso, la capacidad predictiva de árbol de decisión era muy pobre porque el número de registros era muy pequeña.

2.2.1.4 Una Herramienta para la Predicción de Riesgos en Proyectos de Software usando Modelos en Redes Bayesianas.

Según Reyes (2006) el desarrollo de sistemas de software es un proceso costoso, difícil y con altas tasas de fracaso. Siempre está presente la posibilidad de ocurrencia de ciertos eventos que se traducen en complicaciones para su realización a los cuales se les denomina riesgos.

Reyes (2006) indica que los factores que afectan el éxito o fracaso de un proyecto, pueden ser clasificados como riesgos, factores críticos de éxito, y mitigantes. Los riesgos pueden incluir eventos del proceso de desarrollo de software que amenazan el éxito del proyecto. Los factores críticos de éxito son un grupo de factores que el equipo de desarrollo debe asegurarse que estén presentes, ya que con su ausencia existe una alta

posibilidad de que el proyecto fracase. Los mitigantes son acciones o actividades que un equipo de desarrollo puede realizar cuando un riesgo requiere ser eliminado o mitigado.

Reyes (2006) en su estudio indica que se confeccionaron encuestas repartidas a nivel nacional en empresas de desarrollo de software y tecnologías de información chilenas. La información obtenida en las encuestas, junto con la clasificación de éxito/fracaso dada a los proyectos, permitió definir modelos probabilísticos para la predicción del nivel de riesgos de un proyecto de software.

La recolección de datos le permitió la creación de modelos probabilísticos capaces de predecir el nivel de riesgo de proyectos de software en desarrollo; para ello utilizó redes bayesianas con la ayuda del software HUGIN.

Las redes bayesianas, son una técnica de minería de datos, que le permitieron representar el conocimiento experto y fue una excelente herramienta en la representación de la incertidumbre frente a ciertos eventos.

2.2.2 Antecedentes temáticos en Colombia

2.2.2.1 Estudio de Factores Críticos de Éxito Local e Internacional para Empresas de la Industria del Software.

Según Merchan (2007). Desde hace años existe preocupación a nivel mundial de por qué son exitosos los proyectos de tecnología. Actualmente se puede considerar el éxito como un equilibrio entre el tiempo, costo, desempeño, alcance y satisfacción.

En el trabajo realizado por Merchan, se adelantó una investigación sobre los factores (planeación, alcance, recompensas y producto final) que influyen en el éxito de los proyectos de la red de empresas de software – Parquesoft; realizando una comparación entre estas, empresas externas y estudios internacionales.

En cuanto al factor planeación Merchan (2007) observó que para las empresas externas a Parquesoft el uso de software para la administración de proyectos y definir la asignación de la administración de proyectos son las variables más importantes, en cambio para las empresas de Parquesoft definir la administración de calidad y definir las medidas para el éxito de los proyectos son las variables más importantes. Variables como la comunicación entre los administradores y la organización y la participación de los administradores de proyecto desde la fase inicial son variables comunes y tienen una importancia similar en ambos casos. Comparados con el nivel internacional encontró que tanto en las empresas de Parquesoft, como a nivel internacional la variable más importante es la participación de los administradores de proyecto durante la fase inicial.

En cuanto al factor gestión del alcance el autor indicó que para las empresas del parque, las empresas externas y a nivel internacional los objetivos del proyecto son lo más importante en la fase de definición. Observó igualmente que las empresas del parque y las empresas externas también consideran importante el plan de cronograma del proyecto y la comunicación. Identificó que para Parquesoft el plan de cronograma del proyecto es

lo más importante mientras que para las empresas externas la actividad más importante son los objetivos y a nivel internacional es más importante la comunicación

En cuanto al factor de recompensas el autor evidenció que para las empresas externas realizar una evaluación anual de rendimiento, implementar nuevas tecnologías, la entrega de elogios en público y construir motivación dentro de la organización son variables muy importantes que conllevan al éxito de los proyectos. De igual forma encontró que para Parquesoft y las empresas externas las variables flexibilidad de horario e implementar nuevas tecnologías son variables comunes muy importantes para alcanzar el éxito de los proyectos.

Para Parquesoft encontró como variables relevantes los horarios flexibles, la implementación de nuevas tecnologías y la seguridad laboral.

Finalmente el autor en cuanto a la definición de éxito en el sistema final indica que son variables comunes que los requerimientos fueran cubiertos por el sistema terminado, el sistema funcione como fue planeado y el proyecto tenga una vida útil importante. La percepción que se tiene del sistema final difiere bastante. A nivel internacional se considera más importante que los proyectos aporten nuevos conocimientos y experiencias a los usuarios finales que lo utilizan, mientras que para las empresas del parque lo más importante es que los usuarios finales mejoren los procesos, tareas y/o actividades que realizan dentro de sus empresas, logrando un impacto positivo y la utilización del sistema final.

En Parquesoft se puede ver un impacto positivo frente al nivel internacional, porque ellos consideran vital que los proyectos se terminen dentro del presupuesto y cronograma,

razones que hacen exitosos los proyectos. Además el parque se preocupa por la satisfacción del cliente y por el sistema final (cumplir con lo requerido, dentro de costos asequibles).

2.2.3 Antecedentes metodológicos en Colombia

2.2.3.1 Using Data Mining to Describe Long Hospital Stays.

En un trabajo realizado por Gómez & Abásolo (2008) indicaron que mediante la detección de patrones desconocidos a partir de datos existentes, la minería de datos puede proporcionar nuevos conocimientos para resolver un problema en el área de salud. Sin embargo, hay varios desafíos para la aplicación de tales técnicas de descubrimiento de conocimientos en un contexto de salud. Entre ellas se encuentra el hecho de que las fuentes de datos médicos son con frecuencia heterogéneas y no estructuradas y que la complejidad de los conocimientos médicos es difícil de adoptar para un experto en la minería de datos sin formación médica.

De esta forma plantearon un proyecto de minería de datos para ayudar a describir y entender las estancias largas de pacientes en los hospitales en el hospital universitario de Colombia.

El proyecto planteado por los autores sigue los pasos de la metodología CRISP-DM. CRISP-DM (Cross-Industry Standard Process for Data Mining) en el cual indicaron que es una metodología del proceso de minería de datos creada por un grupo de

organizaciones de diferentes industrias con gran experiencia en este tipo de descubrimiento de conocimiento. Esta metodología se adapta a diferentes contextos, ya que es independiente de la industria y de la herramienta. El enfoque recomendado por esta metodología incluye plantear el problema en términos de negocio y traducirlo a un problema de minería de datos, la elección de un algoritmo de minería de datos, perfilado y limpieza de datos, correr varios modelos de minería de datos, análisis de los resultados con la ayuda de expertos en negocios y hacer varias iteraciones del proceso hasta que se obtienen los resultados adecuados. Finalmente presentaron en su trabajo tareas y hallazgos interesantes encontrados en cada una de las fases de CRISP-DM del proyecto.

2.2.3.2 CMIN - herramienta case basada en CRISP-DM para el soporte de proyectos de minería de datos.

Cobos, Zúñiga, Guarín, León & Mendoza. (2010). Indicaron que existen varias metodologías para orientar el proceso de minería de datos. Entre esas metodologías se destacan CRISP-DM y SEMMA (Sample, Explore, Modify, Model, Asses). SEMMA se centra en las características técnicas del desarrollo del proceso, mientras que CRISP-DM mantiene como foco central los objetivos empresariales del proyecto. Debido a ello, CRISP-DM comienza realizando un análisis del problema empresarial para su transformación en un problema técnico de minería de datos. CRISP-DM define una estructura para proyectos de minería de datos y suministra la orientación para su ejecución.

En la herramienta CMIN presentada por los autores especificaron un modelo conceptual formado por diferentes módulos en el que se destaca el módulo de proceso, ya que es en

este donde se hace énfasis en CRISP-DM, según los autores la definición de procesos representa la acción de registrar un proceso mediante la agregación y definición de sus pasos, campos o actividades que se proponen para el desarrollo de un proyecto de minería de datos. El módulo de gestión de procesos permite definir nuevos procesos de minería de datos. A continuación se explica de modo general la forma como se registra CRISP-DM en la herramienta CMIN. Primero el usuario registra la información básica del proceso (nombre, estado y descripción) y luego define los pasos y campos del proceso. En cada paso se define su nombre, el tipo de paso en la jerarquía del proceso, una descripción y el conjunto de campos (información que el desarrollador del proyecto de minería de datos deberá registrar en ese paso).

El usuario también indica el resultado de la edición de los pasos del proceso CRISP-DM en CMIN. Después se lleva a cabo la edición de los campos del paso. Posteriormente se presenta un formulario en el que se le solicita al editor o experto en minería el registro de los campos para cada paso. En cada campo se debe incluir una descripción, el tipo de campo, que define si es una actividad o sugerencia, y si utiliza *workflow*.

2.2.3.3 Sistema de apoyo para la acreditación de la calidad de programas académicos de la universidad de Caldas aplicando técnicas de minería de datos.

Según González Cardona (2011). Una de las principales dificultades que enfrenta el sistema educativo actual es la deserción, su valor acumulado llega a niveles del 45% a

nivel nacional y del 33% en la Universidad de Caldas; una de sus principales causas es la deserción de tipo académico, por lo que se hace primordial la definición de un indicador que logre medir el rendimiento académico en sus diferentes dimensiones (la excelencia, la eficiencia y la eficacia). Una vez definido el indicador, es fundamental determinar los factores que inciden en este con el objeto de tomar acciones que tengan el mayor impacto posible; estos factores incluyen elementos de identidad, socioeconómicos, vocacionales, de estudios previos, del entorno familiar entre otros y sus relaciones, por lo que un análisis multivariado debe ser el tipo de modelo que los caracterice, la regresión logística y su amplio uso en la determinación de factores de riesgo o de protección fue el análisis seleccionado, tanto por esta característica como por su manejo de variables de tipo numérico como categóricas. Este proceso de extracción de conocimiento a partir de los datos KDD, que ha estado en auge en los últimos años en los ambientes educativos, se desarrolló utilizando la metodología CRISP-DM presentando para cada fase las parametrizaciones y resultados obtenidos; comprendiendo desde la extracción de información de la base de datos del Sistema de Información Académica SIA, su transformación, validación, el cálculo de los índices y el indicador de rendimiento académico IAR, el análisis de Regresión Logística por programa y nivel de avance en créditos del programa, hasta la generación de informes de tipo descriptivo como del modelo. Los resultados finales indicaron a González Cardona los factores de riesgo y de protección en el rendimiento académico para los estudiantes de cada programa presencial en diferentes momentos de su paso por la universidad.

2.3. Justificación

La empresa de *Outsourcing* de TI para la cual se está realizando este trabajo, para el presente año tiene como objetivo ampliar aún más su rango de alcance procurando dar soporte a las necesidades de negocio de nuevos grandes clientes, acarreado consigo la contratación de personal que soporten tal iniciativa.

Para dar cumplimiento y apoyar este objetivo, mejorar el éxito de sus proyectos con sus antiguos clientes como los futuros es un punto de gran interés en el que se requiere identificar profundamente los factores de éxito de sus desarrollos, que ante los posibles ambientes hostiles y competencia a los que se enfrentará pueda garantizar un producto final excelente; no solo aprobado por sus usuarios finales sino también por el grupo de trabajo que está velando por los buenos resultados.

Pese a la alta experiencia de la compañía en el desarrollo de proyectos de software, ha tenido sus tropiezos que han afectado su progreso y relaciones con los clientes. La búsqueda constante para garantizar mejores resultados en sus productos y generar satisfacción con sus clientes es una tarea sin fin.

Se han implementado diferentes metodologías, marcos de trabajo y propuestas que den el punto clave para llevar la balanza con gran peso hacia los aspectos positivos y exitosos; sin embargo obtener una efectividad y una regla definitiva que determinen el buen camino a seguir es una tarea que a diario permanece inconclusa.

De aquí que aprovechar el conocimiento como fuente valiosa de información toma fuerza para desarrollar propuestas sobre la problemática; se pretenden entonces mediante el uso de herramientas y la aplicación de técnicas de minería de datos obtener información relevante que permitan generar patrones que soporten la toma de decisiones en la empresa y ayude a fortalecer la consecución de proyectos exitosos, tomando como exitosos proyectos que se ejecutan dentro del tiempo, alcance y presupuesto esperado.

2.4. Formulación del problema

¿Qué factores influyen en la consecución de proyectos exitosos en la empresa de *outsourcing* de TI de la ciudad de Medellín?

2.5. Objetivo

2.5.1. Objetivo General

Descubrir conocimiento en una empresa de *outsourcing* de TI de la ciudad de Medellín aplicando técnicas de minería de datos que permita identificar potencialidades en el éxito de los proyectos de desarrollo de software.

2.5.2. Objetivos Específicos

- Realizar una búsqueda bibliográfica sobre los factores que influyen en la inyección de defectos de desarrollo de software.
- Identificar factores que permitan priorizar proyectos exitosos de desarrollo de software de la empresa de *outsourcing* de TI de la ciudad de Medellín.
- Realizar una descripción e interpretación de los datos de los proyectos de desarrollo de software objetos de estudio.
- Construir un modelo multivariado que permita predecir el éxito o fracaso de un proyecto de desarrollo de software en la empresa de *outsourcing* de TI de la ciudad de Medellín.
- Realizar una simulación del modelo multivariado con los datos de los proyectos de desarrollo de software objeto de estudio.

2.6. Resultados esperados

1. Documento con el estudio bibliográfico sobre los factores que influyen en la detección de defectos de desarrollo de software.
2. Documento con la descripción e interpretación de los datos de los proyectos de desarrollo de software objetos de estudio.
3. Documento con la identificación de factores que inciden en el éxito de proyectos de software de la empresa de *outsourcing* de TI.
4. Documento con la estructura de análisis estadístico y los modelos definidos que describan los factores que inciden en el éxito de los proyectos de software en la empresa de *outsourcing* de TI.

3. ESTRATEGIA METODOLÓGICA

3.1. Metodología

Para este proyecto de investigación aplicada, la metodología seleccionada fue CRISP-DM, por el éxito presentado en los proyectos de minería de datos y porque que abarca todos los objetivos propuestos.

El ciclo de vida del proyecto según la metodología CRISP-DM está basado en 6 fases cambiantes entre sí, lo cual lo postula como un ciclo en constante movimiento.

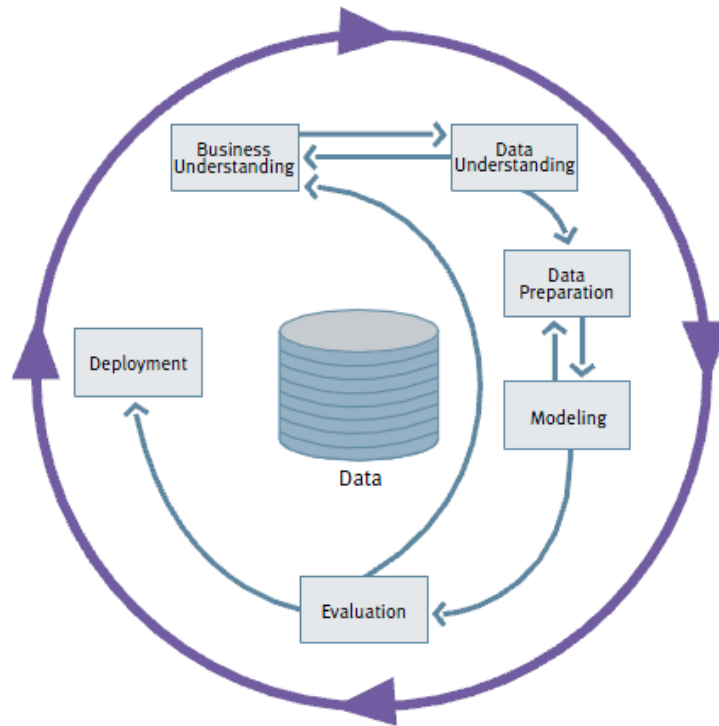
Los datos fueron recolectados partiendo de la identificación de factores que influyen en la inyección de defectos de desarrollo de software, ya que hay un interés de la compañía de evaluar cuáles de estos factores pueden aplicar en los proyectos de la compañía e identificar y analizar cuales influyen en la consecución de proyectos de software exitosos; se realizó una encuesta vía internet de pregunta cerrada a los involucrados en el desarrollo de los proyectos de software basada en los factores identificados en la primera sección.

La población objetivo hacia la cual estuvo orientada la encuesta fue a los líderes técnicos o arquitectos de cada equipo de trabajo, se seleccionó esta persona ya que se considera que es la persona que tiene el conocimiento tanto técnico como de gestión de los proyectos que han culminado, adicionalmente, aunque hay mucha rotación de personal en la compañía, el arquitecto es uno de los cargos que se conservan con más tiempo y por lo

tanto da la opción de que se tenga mayor cantidad de información de los proyectos terminados y así poder brindar una información sustancial y valedera, ya que la primicia para un proyecto de minería de datos parte de la cantidad de datos y consistencia de información que se tenga; la idea finalmente es poder evaluar todos los proyectos que hayan sido culminados y de los cuales se pueda obtener la información requerida por medio de encuestas.

El total de proyectos culminados de los cuales se tuvo información para evaluar por medio de las encuestas fue de 83, ya que se hizo un censo de los líderes o arquitectos actuales de la compañía y la cantidad de proyectos sobre los que cada uno ha tenido participación directa y que haya culminado su desarrollo.

Figura 5 Metodología CRISP-DM



Nota fuente: Tomado de IBM Corporation (2010) (p. 5)

3.1.1. Comprensión del negocio.

Según IBM Corporation (2010) esta fase inicial se centra en la comprensión de los objetivos y requisitos desde una perspectiva empresarial del proyecto, luego convertir este conocimiento en la definición de un problema de minería de datos y en un plan preliminar diseñado para alcanzar los objetivos. Las tareas generales a realizar para esta primera fase son:

- A. Determinar los objetivos del negocio.

- B. Evaluar la situación.
- C. Determinar los objetivos de la minería de datos.

3.1.2. Comprensión de datos.

Según IBM Corporation (2010) la fase de comprensión de datos se inicia con la recolección de datos y continúa con las actividades que le permiten familiarizarse con los datos, identificar los problemas de calidad de datos, y detectar subconjuntos interesantes para formar hipótesis sobre la información oculta de los datos iniciales. Las tareas generales a realizar para esta segunda fase son:

- A. Recolección inicial de datos.
- B. Descripción de datos.
- C. Exploración de datos.
- D. Verificar la calidad de los datos.

3.1.3. Preparación de datos.

Según IBM Corporation (2010) la fase de preparación de datos cubre todas las actividades necesarias para construir el conjunto de datos final (datos que se introducen en las herramientas de modelado) de los datos iniciales sin procesar. Las tareas de preparación de datos son susceptibles a ser realizadas muchas veces y no en un orden

prescripto. Las tareas incluyen la selección de tablas, registros, y atributos, así como la transformación y la limpieza de datos para las herramientas que modelan. Las tareas generales a realizar para esta tercera fase son:

- A. Selección de datos
- B. Limpieza de datos.
- C. Construcción de datos.
- D. Integración de datos.

3.1.4. Modelado.

Según IBM Corporation (2010) en esta fase, varias técnicas de modelado son seleccionadas y aplicadas, y sus parámetros son calibrados a valores óptimos. Por lo general, existen varias técnicas para el mismo tipo de problema de minería de datos. Algunas técnicas tienen requerimientos específicos en la forma de datos. Por lo tanto, volver a la fase de preparación de datos es a menudo necesario. Las tareas generales a realizar para esta cuarta fase son:

- A. Seleccionar la técnica de modelado.
- B. Generar el diseño de las pruebas.
- C. Construcción del modelo.
- D. Evaluación del modelo.

3.1.5. Evaluación.

Según IBM Corporation (2010) en esta etapa del proyecto, se ha construido un modelo (o modelos) que parecen tener alta calidad desde la perspectiva de análisis de datos. Antes del proceder al despliegue final del modelo, es importante evaluarlo a fondo y revisar los pasos ejecutados para crearlo, para asegurar que el modelo logra adecuadamente los objetivos de negocio. Un objetivo clave es determinar si hay algún problema importante de negocio que no ha sido suficientemente considerado. Al final de esta fase, deberá tomarse una decisión sobre el uso de los resultados de minería de datos. Las tareas generales a realizar para esta quinta fase son:

A. Evaluar los resultados

3.1.6. Despliegue.

Según IBM Corporation (2010) la creación del modelo no es generalmente el final del proyecto. Aunque el objetivo del modelo es de aumentar el conocimiento de los datos, tendrá que ser organizado y presentado de manera que el cliente pueda utilizarlo. Ello a menudo implica la aplicación de modelos "en vivo" en los procesos de toma de decisiones de la organización, por ejemplo, la personalización en tiempo real de páginas Web. Dependiendo de los requisitos, la fase de despliegue puede ser tan simple como la generación de un informe o tan compleja como la realización repetida de un proceso de

minería de datos en toda la empresa. En muchos casos, es el cliente, no el analista de datos, el que lleva a cabo los pasos de despliegue. Sin embargo, incluso si el analista realiza el esfuerzo de despliegue, es importante que el cliente entienda anticipadamente que acciones necesita ejecutar con el fin de hacer uso de los modelos creados. Las tareas generales a realizar para esta sexta fase son:

A. Generar informe final

3.2. Pruebas

Las pruebas del proyecto están especificadas en la tarea general “Generar el diseño de las pruebas”, en la fase de modelado de la metodología CRISP-DM.

3.3. Presupuesto

Para el desarrollo del presente proyecto se va a hacer uso de herramientas de software libre y versiones trial y será desarrollado por el autor con el acompañamiento del asesor; sin embargo se tienen algunos costos asociados:

Costos	Valor
Minutos telefonía celular	80000
Pasajes (viaje reunión con asesor)	150000
Internet	40000
Socializar resultados en la Universidad	150000
Gran total	420000

4. DESARROLLO

4.1. Referente teórico

Día a día las diferentes organizaciones se enfrentan a un crecimiento constante de información que se encuentra almacenada dentro de sus bases de datos digitales, gran parte de estas bases de datos se convierten en repositorios históricos de datos, esta información histórica es útil para explicar el pasado, entender el presente y predecir el futuro, siendo inminente la necesidad de las organizaciones de analizar los datos con el fin de obtener información importante que ayude a tomar decisiones y generar nuevo conocimiento.

4.1.1. Dato, información y conocimiento.

Al tratarse este trabajo sobre generación del conocimiento vale la pena diferenciar los diferentes términos que abarcan este tópico.

Según Montoro, sf. Citado por Fernández & Ponjuante, (2008). Los datos se identifican como “la materia prima de la información, son hechos físicos que no contienen un significado inherente, no incluyen necesariamente interpretaciones u opiniones, y no llevan asociado ningún rasgo indicativo que pueda develar su importancia o su relevancia”. De igual forma sostiene que “la información se identifica en este contexto con el dato dotado de significado. La información debe entenderse como el subconjunto

de datos que adquieren significado para su receptor. Un dato pasa a ser información cuando adquiere significación para su receptor, un dato es o no información en dependencia de si es o no significativo para ese receptor...” de esta forma propone que “la información debe ser identificada, no como un subconjunto especial de datos, sino como el contenido semántico de los datos”. Finalmente respecto al conocimiento el autor indica que este debe identificarse como “la información que se asimila por un individuo y que le permite a éste tomar decisiones y actuar. En este sentido, el conocimiento se encuentra mucho más relacionado con la acción que los datos o la propia información”.

4.1.2. Minería de datos.

Según (Witten & Frank, 2000) citado por Hernández, Ramírez & Ferri (2004), se define la minería de datos como el proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de almacenes de datos en distintos formatos, con el fin de generar patrones que permitan la toma de decisiones en la organización y por lo tanto generar beneficios a la misma; en este sentido la minería de datos se enfrenta a dos grandes retos, el primero poder trabajar con grandes volúmenes de datos procedentes de diferentes almacenes de datos con los problemas que esto genera (ausencia de datos, ruido, intratabilidad... etc....) y como segundo reto el uso de técnicas adecuadas que permita analizar estos datos y extraer conocimiento útil y novedoso.

Existe un término que según Hernández, Ramírez & Ferri (2004), algunos autores lo usan como sinónimo minería de datos y es el “descubrimiento de conocimiento en bases de

datos” (Knowledge Discovery in Databases, KDD), aunque en muchas ocasiones se han utilizado indistintamente ambos términos, existe una diferencia entre ambos, ya que KDD se ha usado para referirse al proceso que consta de una serie de fases, mientras que la minería es una de estas fases.

4.1.3. Técnicas de minería de datos.

Con el fin de llevar a cabo la tarea de minería de datos existen ciertas técnicas que permiten la extracción conocimiento, aplicables en diferentes contextos según las necesidades que se presenten en cada problema, entre las técnicas más utilizadas se encuentran:

- Métodos bayesianos: Se trata de un modelo que permite un doble uso: descriptivo y predictivo; en cuanto al modelo descriptivo, los algoritmos de redes bayesianas se basan en el descubrimiento de relaciones de independencia y/o relevancia entre sus variables; el modelo predictivo usa las redes bayesianas como clasificadores (Hernández, Ramírez & Ferri, 2004).
- Árboles de decisión: son un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que la decisión final a tomar se puede determinar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta alguna de sus hojas; una de sus grandes ventajas es que en su forma más general, las opciones posibles a partir de una determinada condición son excluyentes. Esto

permite analizar una situación, y siguiendo al árbol apropiadamente, llegan a una sola acción o decisión a tomar. (Hernández, Ramírez & Ferri, 2004).

- Redes neuronales: Son un método de aprendizaje cuya finalidad inicial era la de emular los procesadores biológicos de información. Las redes neuronales parten de la presunción de que la capacidad humana de procesar información se debe a la naturaleza biológica del cerebro.

Pueden utilizarse para el reconocimiento de patrones, la comprensión de información y la reducción de la dimensionalidad, el agrupamiento, la clasificación, la visualización etc. (Hernández, Ramírez & Ferri, 2004).

- Algoritmos genéticos: Imitan la evolución de las especies mediante mutaciones, reproducciones y selectividad de los más fuertes. De esta forma los algoritmos genéticos son ampliamente utilizados en procesos de optimización debido a que su principal objetivo es la búsqueda de la efectividad y la eficacia. (Hernández, Ramírez & Ferri, 2004).

- Métodos multivariados:

Según (Montgomery & Runger, 2003).citado por Soto (2009) el campo de la estadística tiene que ver con la recopilación, presentación, análisis y uso de datos para tomar decisiones y resolver problemas; este campo ha desarrollado técnicas univariadas donde se analiza cada una de las variables estudiadas por separado, es decir, se examina solo una variables y las técnicas multivariadas examina más de dos variables, en ese sentido, las técnicas estadísticas para el análisis, también se consideran parte de la minería de datos, debido a que persiguen el mismo objetivo.

Clúster: Agrupan datos dentro de un conjunto de clases partiendo de un criterio de distancia o similitud dentro de cada clase. Su utilización ha proporcionado significativos resultados en lo que respecta a los clasificadores o reconocedores de patrones, como en el modelado de sistemas. Un problema relacionado con el análisis de clúster es la selección de factores en tareas de clasificación, debido a que no todas las variables tienen la misma importancia a la hora de agrupar los objetos.

Análisis factorial: Es una técnica de análisis multivariado que se utiliza para estudiar las correlaciones entre un grupo de variables, es una técnica de reducción de datos que sirve para encontrar grupos homogéneos de variables a partir de un conjunto numeroso de variables, su fin consiste en encontrar el número mínimo de variables capaces de explicar el máximo de información contenida en los datos. (de la Fuente, 2011).

Analiza la estructura de las relaciones entre un gran número de variables, utilizando esta información se calcula un conjunto de dimensiones latentes, conocidas como factores, que buscan explicar dichas relaciones. Es, por lo tanto, una técnica de reducción de datos dado que si se cumplen sus hipótesis, la información contenida en la matriz de datos puede expresarse, en un número menor de dimensiones representadas por dichos factores.

La correlación es una técnica estadística usada para determinar la relación de dos o más variables; El Coeficiente de Correlación es un valor cuantitativo de la relación entre dos o más variables el cual puede variar desde -1.00 hasta 1.00, un valor de 1 indica relación lineal perfecta positiva; un valor de -1 indica una relación lineal perfecta negativa; no existe relación entre las variables cuando el coeficiente es de 0.00. (Ramón, sa).

Dentro del análisis factorial se encuentra el análisis de componentes principales el cual se utiliza para descubrir la estructura latente (dimensiones) de un conjunto de variables. Reduce el espacio de atributos a partir de un número mayor de variables a un número menor de las mismas y, como tal, es un procedimiento "no dependiente" (es decir, no se asume una variable dependiente específica). (Terradez Gurrea, sa).

Regresión logística multivariado: La regresión logística es un tipo especial de regresión que se utiliza para predecir y explicar una variable categórica binaria, ejemplo: Éxito o fracaso, por esto, la variable de salida o explicada se convierte en una variable binaria con valores de cero (0) o uno (1), para representar un fracaso o un éxito. Sea Y la variable de respuesta dicotómica y explicada por una

combinación lineal de variables $X_1; \dots; X_n$, se puede considerar que $Y = 1$ si el evento ocurre, o $Y = 0$ si no lo hace. En este modelo, el problema es encontrar la estimación de la probabilidad de ocurrencia de un evento, dado que ha ocurrido x es $p(y = 1|x)$ se indica esta probabilidad está dada por $p(x)$; la probabilidad contraria de que no ocurra el evento dado x es $p(y = 0|x) = 1 - p(x)$. (Soto Jaramillo, 2009)

La aplicación directa del modelo de regresión lineal multivariado:

$$Y_i = \beta_0 + \beta_1 X_{i,1} + \beta_2 X_{i,2} + \dots + \beta_{p-1} X_{i,p-1} + \varepsilon_i$$

No es apropiado en el caso de una variable respuesta dicotómica, dado que las estimaciones que se hagan usando un modelo de regresión lineal puede generar valores diferentes a cero, o a uno. Por lo anterior, es necesaria una transformación lo que sugiera un nuevo modelo el cual se plantea de la siguiente manera:

$$\ln\left[\frac{p(\mathbf{x})}{1 - p(\mathbf{x})}\right] = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p = \beta_0 + \beta' \mathbf{x}$$

Donde:

$$\mathbf{X}_i = \begin{bmatrix} 1 & X_{i,1} & X_{i,2} & \dots & X_{i,p-1} \end{bmatrix}_{1 \times p}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p-1} \end{bmatrix}_{p \times 1}$$

El modelo equivale a suponer las siguientes probabilidades para la ocurrencia del evento y su no ocurrencia contrario, ambas en función de x :

$$p(\mathbf{x}) = \frac{e^{\beta_0 + \beta' \mathbf{x}}}{1 + e^{\beta_0 + \beta' \mathbf{x}}}, \quad 1 - p(\mathbf{x}) = \frac{1}{1 + e^{\beta_0 + \beta' \mathbf{x}}}.$$

González Cardona (2011).

El ajuste del modelo de regresión logística consiste en encontrar los estimadores más apropiados de los parámetros o coeficientes β_i , utilizando los datos observados. Existen varios procedimientos numéricos que permiten estimar los parámetros β_i . El método de mínimos cuadrados utilizado en la regresión lineal no se puede utilizar para estimar los parámetros β_i en regresión logística. En este caso, se requiere de procedimientos numéricos más complejos, un método usado con frecuencia es el algoritmo de Marquardt.

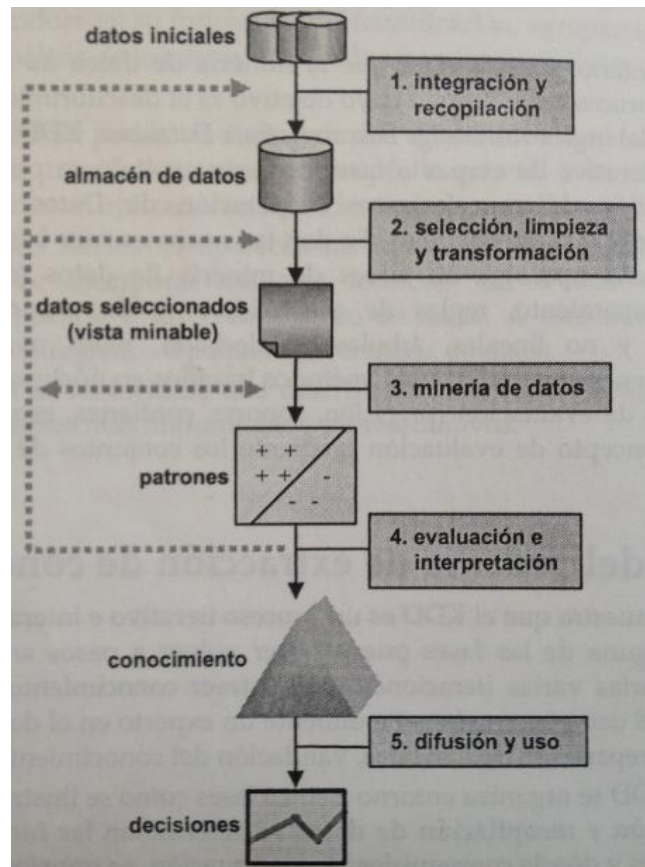
Después de construir un modelo de regresión logística, se debe validar la bondad de ajuste con el fin de poderlo utilizar para describir o predecir valores no sólo futuros, sino para otros valores en las variables independientes no observados o medidos. Se verifica la significancia estadística de las variables explicativas, de manera global o particular, y el grado de cumplimiento de los supuestos impuestos al modelo de regresión logística. (Soto Jaramillo, 2009).

4.1.4. Descubrimiento de conocimiento en bases de datos.

Fayyad et al, 1996^a, citados por Hernández, Ramírez & Ferri (2004), definieron KDD como “El proceso no trivial de identificar patrones válidos, novedosos potencialmente útiles y, en última instancia, comprensibles a partir de los datos”.

A continuación se presenta el proceso de KDD:

Figura 6 Fases del proceso KDD



Nota fuente: Tomado de Hernández Orallo, J. Ramírez Quintana, M.J & Ferri Ramírez, C. (2004) (p. 93)

Este esquema del proceso permite clarificar la diferenciación con minería de datos. KDD es el proceso global de descubrir conocimiento desde las bases de datos mientras que la minería de datos se refiere a la aplicación de los métodos de aprendizaje y estadístico para la obtención de patrones y modelos.

Este proceso de KDD consta de una secuencia iterativa de fases:

Fase de integración y recopilación de datos: Se determinan las fuentes de información que pueden ser útiles y donde conseguirlas.

Fase de selección, limpieza y transformación: seleccionar o segmentar los datos de acuerdo a su relevancia con base en los objetivos planteados para la extracción de conocimiento; limpieza de la información, eliminando información innecesaria, consolidando información y unificando formatos que provenga de múltiples fuentes; transformar la información recolectada con el objeto de que pueda ser utilizable y navegable preparándola para la etapa de minería de datos.

Fase de minería de datos: se decide cual es la tarea a realizar (clasificar, agrupar etc.) y se elige el método que se va a utilizar.

Fase de evaluación e interpretación: Se evalúan los patrones y se analizan por los expertos, y si es necesario se vuelve a fases anteriores para una nueva iteración.

Fase de difusión: Se hace uso del nuevo conocimiento y se hace partícipe de él a todos los posibles usuarios.

4.2. Factores que influyen en la inyección de defectos de desarrollo de software.

Los proyectos de software no están exentos de que se le inyecten defectos en cualquiera de sus fases de desarrollo, desde la captura de requisitos hasta su

implementación y despliegue, la incidencia de estos defectos dependen de diferentes factores como son los humanos, los técnicos, entre otros; dependiendo de las características de los proyectos de software, estos pueden tener muchos o pocos defectos, diferentes o similares; y a pesar de que un factor pueda ser considerado como un elemento para la inyección de defectos en el desarrollo de software, también puede intervenir en la consecución de proyectos exitosos; es por esto que la compañía se interesó en identificar mediante una investigación bibliográfica que factores se pueden determinar como incidentes en la inyección de defectos, determinar cuáles factores son medibles y si se tiene información de los proyectos de la empresa para finalmente realizar un estudio que permita identificar a partir de estos factores potencialidades en el éxito de los proyectos de software.

Son varios los autores que han investigado estos tipos de factores, entre ellos están:

Zhan & Pham (2000) realizaron una investigación empírica de 13 empresas que participan en la industria del software para identificar los factores que pueden afectar la fiabilidad y confiabilidad del software. El estudio utilizó la encuesta como instrumento para la recolección de la información, fueron encuestados los principales participantes de los proyectos entre ellos estaban: gerentes, desarrolladores y los tester; utilizaron dos técnicas para analizar la información el método del peso relativo y la técnica de análisis de varianza (ANOVA), finalmente como resultado obtuvieron 32 posibles factores implicados en las diferentes etapas del desarrollo de software:

- Complejidad del programa
- Habilidad de los programadores

- Cobertura de las pruebas
- Esfuerzo en las pruebas
- Entorno de las pruebas
- Frecuencia en los cambios
- Metodología de pruebas
- Análisis de requerimientos
- Porcentaje de código reutilizado
- Relación entre el diseño detallado y los requisitos
- Nivel de tecnologías de programación
- Documentación
- Carga de trabajo del programa
- Herramientas de pruebas
- Organización de los programadores
- Conocimiento del dominio
- Dificultad de programación
- Metodologías de diseño
- Naturaleza humana (Errores)
- Gestión del desarrollo
- Asignación de recursos de pruebas
- Esfuerzo en la programación
- Categorías del programa
- Estándares de trabajo
- Software del sistema

- Volumen de los documentos del diseño del programa
- Tamaño del equipo de desarrollo
- Lenguaje de programación
- Procesador
- Dispositivo de telecomunicaciones
- Dispositivo de entrada/salida
- Dispositivo de almacenamiento

Stutzke & Smidts (2001) realizaron un modelo estocástico de introducción y retiro de defectos durante el desarrollo del software, los autores indicaron que los modelos de fiabilidad del software estiman la intensidad de fallas durante las pruebas del software, mientras que estos modelos usan herramientas ingenieriles, estos no pueden ser aplicados a las fases tempranas del ciclo de vida (requerimientos, análisis, diseño, implementación) y no proporcionan información de los errores humanos que se producen durante el desarrollo del software, de igual forma indican que hay dos categorías de errores humanos que se producen durante el proceso del desarrollo del software la primera consta de los errores realizados durante la fase de requisitos, análisis, diseño e implementación y la segunda la depuración de los errores realizados durante los intentos de eliminar las fallas identificadas durante la inspección y pruebas del software.

El modelo presentado por los autores relaciona la función de intensidad del fallo en el desarrollo y depuración de errores en todas las fases del ciclo de vida del software, usaron la metodología de probabilidad de éxito (*SLIM*) por sus siglas en ingles. Con el fin de calcular el *SLI* de ocurrencia de errores en el desarrollo del software, tuvieron un

acompañamiento de un equipo de desarrollo que les ayudaron a determinar y clasificar las influencias más importantes, teniendo así los siguientes factores:

- Experiencia
- Capacidad
- Complejidad del software
- Presión del calendario
- Uso de métodos
- Comunicación
- Relación del equipo
- Estilo de la gestión
- Métodos de integración de procesos
- Volatilidad de los requerimientos

Jacobs et al (2004) realizaron un trabajo con el fin de explorar las causas de los defectos en productos desarrollados por equipos virtuales, el objetivo de la investigación era crear un marco de trabajo que permitiera la reducción de defectos residuales en productos de software desarrollado por equipos virtuales. Con el fin de identificar las causas de los defectos los autores realizaron diferentes reuniones con equipos expertos de trabajo, estos equipos estaban conformados por personas que cumplían las siguientes características:

- Más de 3 años de experiencia práctica en el área de desarrollo virtual.
- Participación directa en pruebas de los productos, soporte técnico, o en la gestión de proyectos.

- Tener una amplia experiencia en el análisis causal de defectos en los proyectos de desarrollo virtual.
- Estar directamente involucrado en múltiples proyectos de desarrollo virtual en los últimos 2 años.

Finalmente como resultado de la investigación obtuvieron las siguientes causas de la inyección de defectos en proyectos de equipos virtuales, de acuerdo a una categorización de riesgos del proyecto:

Categoría del riesgo	Causas de la inyección de defectos
Comunicación	Especificación de requerimientos técnicos Diseño global
Proceso	Diseño de la arquitectura del sistema Especificación de requerimientos técnicos Diseño global Implementación Pruebas de integración
Organización	Requisitos técnicos del sistema Especificación de requerimientos técnicos Implementación Pruebas de integración
Tecnología	Diseño de la arquitectura del sistema Especificación de requerimientos técnicos Implementación

Leszak, Perry & Stoll (2002) reportaron un análisis retrospectivo de las solicitudes de modificación de defectos (MRS) descubiertos mientras realizaban la construcción, pruebas y despliegue de una versión de un elemento de red como parte de una red de transmisión óptica. El estudio consistía en 3 investigaciones: un estudio de análisis de defectos causa-raíz (RSA), un estudio métrico del proceso, y una investigación de la complejidad del código. Según los autores los 3 trabajos difirieron en las cantidades relacionadas de los defectos encontrados, pero todos tenían en común el objetivo de identificar indicadores tempranos de calidad.

La misión del estudio RCA era:

- Analizar la muestra de defectos MR.
- Encontrar las causas raíces sistemáticas de los defectos.
- Analizar las principales MRs notificados por el usuario del elemento de red.

Proponer acciones de mejora, como insumo para proyectos de desarrollo en curso, a fin de reducir el número de defectos críticos (gravedad 1 y 2 MRS) y para reducir el costo del re trabajo.

Finalmente los autores realizaron una clasificación de los tipos de defectos, así:

Implementación	Interface	Externos
Diseño/uso de datos	Diseño/uso de datos	Entorno de desarrollo
Asignación/uso de recursos	Diseño/uso de la funcionalidad	Entorno de pruebas
Manejo de excepciones	Protocolos de comunicación	Trabajo concurrente
Algoritmo	Coordinación de procesos	Otros
Funcionalidad	Interacciones inesperadas	
Rendimiento	Cambios de coordinación	
Dificultades en el lenguaje	otros	
Otro		

Jacobs, Moll, Kusters, Trienekens & Brombacher (2007) realizaron un estudio que consistía en la identificación de los factores que influyen en la inyección y detección de defectos. El estudio de los autores es una parte de un proyecto de investigación más amplio, con el objetivo de reducir el número de defectos residuales en productos intensivos en software, mediante el uso de los factores que influyen en la disminución de la inyección de defectos y en el aumento de la detección de los defectos. Como primera instancia, realizaron una extensa búsqueda en la literatura para encontrar los factores que influyen y procesaron los factores para lograr conjuntos consistente de factores y sin duplicaciones; Como segundo paso , realizaron un análisis de conglomerados para reducir el número de factores influyentes y tener conjuntos gestionables para su

aplicación práctica; como último paso, los grupos finales de los factores se obtuvieron mediante la interpretación de expertos de los resultados del análisis de *clúster*.

El agrupamiento y los factores que inciden en la inyección de defectos identificados por los autores son:

A. Capacidad de los desarrolladores

- Conocimiento de los desarrolladores acerca de los procesos de desarrollo.
- Nivel de experiencia con la plataforma o entorno del proyecto.
- Experiencia con el entorno de desarrollo.
- Nivel de experiencia con herramientas y lenguajes de desarrollo.
- Conocimiento de los desarrolladores acerca de las herramientas utilizadas para apoyar las actividades de desarrollo.
- Capacidad del personal para dar estimaciones precisas de las actividades de prueba.

B. Conocimiento del dominio

- Nivel en que los desarrolladores entienden la arquitectura y resuelven los riesgos asociados.
- Experiencia previa con el desarrollo de una aplicación similar.
- Conocimiento del desarrollador acerca de los aspectos del negocio del producto a probar.
- Conocimiento de los desarrolladores sobre el dominio para el que está destinado el producto.
- Conocimientos sobre el producto a desarrollar.

- Nivel de experiencia del desarrollador en el entorno del usuario y las operaciones.
- Conocimiento del entorno de aplicación.
- Conocimiento acerca de cómo las partes del producto se han implementado por otras partes.

C. Composición del equipo

- Experiencia en programación.
- Nivel de habilidad de los desarrolladores.
- Nivel de habilidad de los diseñadores.
- Experiencia del equipo.
- Capacidad del equipo.
- Calidad y habilidad del equipo de desarrollo.
- Dependencia de poco personal clave.
- Disponibilidad de personal de desarrollo adecuado.
- El uso de personal técnico capacitado.
- El uso de especialistas.
- Tamaño del equipo del proyecto.
- Composición del equipo de desarrollo de software complejo.
- Número de equipos que participan en el desarrollo.
- Complejidad de la organización general del proyecto del cual el equipo de desarrollo es sólo una parte.

D. Distribución del Equipo

- Distribución geográfica en locaciones separadas de los desarrolladores.

- Diferencias culturales entre los grupos, sub- proyectos y partes relacionadas.

E. Colaboración

- Grado de estrés mental desarrolladores.
- Nivel en que los desarrolladores y el equipo demuestran flexibilidad en relación con las circunstancias del proyecto.
- Relaciones y cohesión del equipo de desarrollo.
- Continuidad y la estabilidad del equipo.
- Actitud y compromiso de los desarrolladores.
- Grado en que los sub contratistas apoyan el trabajo de desarrollo.
- Grado en que los clientes apoyan el trabajo de desarrollo.
- Grado en que los usuarios finales apoyan el trabajo de desarrollo.

F. Madurez de Gestión Empresarial

- Experiencia de la empresa en la implementación de cambios.
- Experiencia de la empresa en un proyecto de esta naturaleza.
- Experiencia y familiaridad con la aplicación.
- Estabilidad del gerente.
- Estilo de gestión del gerente.
- Apoyo de la dirección.

G. Complejidad del producto

- Número de pasos de integración y su ubicación física.
- Complejidad del problema.
- Número y la complejidad de interfaces externas del producto a desarrollar.

- Nivel de innovación de las características del producto a desarrollar.
- La complejidad del producto a ser desarrollado.
- Tamaño del producto a ser desarrollado.
- Restricciones por tamaño de la memoria principal o disponibilidad de almacenamiento.
- Grado en que el software debe ser diseñado para ser reutilizable en otros productos o proyectos.
- Grado en que el software debe cumplir con su función en condiciones específicas durante un período de tiempo especificado.
- Grado en que el software debe proporcionar las respuestas adecuadas, tiempos de procesamiento y las velocidades de rendimiento.

H. Comunicación

- Lentitud en la comunicación y distribución de cambios.
- Información oculta por otros / sub-proyectos.
- Comunicación entre los desarrolladores.
- Comunicación con el equipo de desarrollo y a los participantes fuera del equipo.
- Errores de comunicación entre los equipos de desarrollo.
- Competencias de la persona de contacto entre un equipo de desarrollo y partes relacionadas.

I. Madurez de Gestión de Proyectos

- Definición de los roles y responsabilidades de las actividades de desarrollo.

- Definición de las responsabilidades y autoridades para actividades de desarrollo.
- Disponibilidad de tiempo suficiente para las actividades de desarrollo.
- Reportes del avance del proyecto.
- Planificación de proyectos.
- Adherencia a los planes.
- Precisión de las métricas utilizadas en la planificación y gestión de proyectos.
- Disponibilidad de datos históricos para apoyar la planificación y gestión de proyectos.
- Suposiciones en la planificación de proyectos, diseño e implementación.
- Grado en que la gestión del proyecto hace su tarea.
- Procesos de desarrollo y prueba alineados.
- Ejecución de tareas de pruebas no planificadas.

J. Perturbaciones externas

- Impacto de los eventos que hacen que el desarrollo se desvíe del plan.

K. Madurez de Procesos

- Grado en que se adquiere el conocimiento, se almacena y se mantiene.
- Madurez del proceso de desarrollo.
- Uso de los procesos de desarrollo.
- Uso de procedimientos disciplinados de control de calidad.
- Utilización de la metodología de diseño disciplinada.
- Utilización de la metodología de análisis de requisitos.

- Uso de procedimientos disciplinados para registrar las modificaciones de especificaciones, preguntas y respuestas de los requisitos y los acuerdos de la interfaz.
- Uso de procedimientos disciplinados para identificar y establecer líneas de base del producto.
- Uso de procedimientos de gestión de configuración disciplinados.
- Trazabilidad entre los requisitos del producto y las pruebas derivadas.
- Calidad del software de terceros y reutilización de código.
- Calidad de los entregables de sub-proyectos.

L. Control de cambios

- Seguimiento y control de los cambios en el producto.
- Grado en que se gestionan los requisitos o las especificaciones de cambios son gestionados.

M. Calidad de la documentación

- Documentación creada a lo largo del ciclo de vida.
- Calidad e integridad de la documentación del sistema.
- Calidad de la documentación.
- Cantidad de los documentos de diseño.

N. Requerimientos

- Pruebas del producto incluido en el diseño.
- Estabilidad de los requisitos.
- Especificación de cambios durante la implementación.

- Cambios o incremento de los requisitos del producto durante y después el diseño.
- Especificación de características innecesarias.
- Malinterpretación de los requisitos.
- Calidad de las especificaciones de requisitos.
- Requisitos adecuados como entrada para la implementación.
- Requisitos permiten múltiples interpretaciones, múltiples supuestos o requisitos implícitos.
- Adecuación de los requisitos para la implementación base.

O. Entorno de desarrollo

- Nivel de acceso a sistema de desarrollo.
- Disponibilidad de los recursos de hardware para el desarrollo
- Estabilidad del hardware y el software que soportan las actividades de desarrollo.
- Utilización de las herramientas de desarrollo de software.
- Eficacia de las herramientas de software.

P. Innovación

- Uso de las nuevas tecnologías.
- Madurez de la tecnología utilizada.

4.3. Comprensión del negocio

4.3.1. Objetivos institucionales.

- ✓ Descubrir factores de éxito de los proyectos de la compañía.
- ✓ Mejorar el rendimiento de los proyectos de software.
- ✓ Obtener ejecuciones futuras de proyectos exitosos.
- ✓ Disminuir la cantidad de ejecución de proyectos no exitosos.

4.3.2. Evaluación de la situación.

- Recursos Humanos: Se cuenta con la asesoría de Msc Javier Hernández Cáceres quien es experto en minería de datos, con la colaboración del líder de arquitectura y procesos de la compañía de quien partió el interés de realizar la investigación y con los diversos líderes técnicos o arquitectos quienes realizaran las encuestas.
- Recursos computacionales: La realización de la encuesta se hizo por medio de la herramienta *form* de *Google Drive*, la cual brindó facilidad para ejecutar la encuesta y suplió las necesidades de la misma.

Para el análisis y procesamiento de la información de las encuestas, se cuenta con un equipo portátil Dell Studio 1557, con procesador *intel core i7* de 1.60GHz y 3Gb de memoria *ram*; adicionalmente, el uso de software y herramientas gratuitas

o versiones trial de las mismas, el único software licenciado con el que cuenta la compañía para soportar la investigación es el *Microsoft Office*.

➤ Datos: después de realizar el estudio bibliográfico de los factores que inciden en la inyección de defectos en el desarrollo de software, y posterior análisis con el líder de arquitectura y procesos de la compañía se seleccionaron los siguientes factores, sobre los cuales él considera serían de mayor interés y viables a evaluar en las encuestas, en algunos casos algunos factores identificados en el estudio bibliográfico se dividieron o renombraron, con el fin de facilitarle el entendimiento por parte de los encuestados y de ampliar la cobertura del factor:

- ✚ Conocimiento de los desarrolladores acerca de los procesos de desarrollo.
- ✚ Grado de experiencia con la plataforma / entorno del proyecto.
- ✚ Grado de experiencia con herramientas y lenguaje de desarrollo.
- ✚ Conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto.
- ✚ Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto.
- ✚ Experiencia previa de los desarrolladores con aplicaciones similares al proyecto.
- ✚ Conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar.
- ✚ Conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto.
- ✚ Conocimientos de los desarrolladores sobre el producto a desarrollar.

- ✚ Capacidad de comunicación: Se dividió en dos:
 - Capacidad de comunicación de los desarrolladores.
 - Capacidad de comunicación de los diseñadores
- ✚ Capacidad de los desarrolladores: se dividió en dos:
 - Capacidad de los desarrolladores para resolver problemas.
 - Capacidad de los diseñadores para resolver problemas.
- ✚ Nivel de habilidad de los desarrolladores: se renombró por:
 - Conocimientos técnicos de los desarrolladores.
- ✚ Nivel de habilidad de los diseñadores: se renombró por:
 - Conocimientos técnicos de los diseñadores.
- ✚ Complejidad de las interfaces externas del producto a desarrollar.
- ✚ Definición de roles y responsabilidades para las actividades de desarrollo del proyecto.
- ✚ Definición de niveles de autoridad para las actividades de desarrollo del proyecto.
- ✚ Adherencia a los planes: se dividió en 2:
 - Cumplimiento con el plan de trabajo respecto al proyecto.
 - Cumplimiento con el plan de trabajo respecto a los procesos de la empresa.
- ✚ Experiencia en programación de los desarrolladores.
- ✚ Experiencia del equipo en proyectos de desarrollo.
- ✚ Uso de personas con postgrado en el proyecto.
- ✚ Tamaño del equipo del proyecto.
- ✚ Nivel de continuidad y estabilidad del equipo del proyecto.

- ✚ Cantidad de interfaces externas del producto (tanto de hardware como software).

- ✚ Tamaño del producto a desarrollar.

➤ Supuestos:

Los líderes técnicos o arquitectos son las personas seleccionadas para responder las encuestas, ya que se consideran tienen todo el conocimiento de la ejecución de los proyectos, tanto en la parte técnica como en la parte de gestión, y en base a esto podrán responder la encuesta de una manera más acertada.

4.3.3. Objetivos de la minera de datos.

- ✓ Identificar y determinar los factores que influyen en la consecución de proyectos de software exitosos.

4.4. Comprensión de los datos

4.4.1. Recolección inicial de datos.

La población objetivo hacia la cual estuvo orientada la encuesta fue a los líderes técnicos o arquitectos de cada equipo de trabajo, se seleccionó esta persona ya que se considera que es la persona que tiene el conocimiento tanto técnico como de gestión de los proyectos que han culminado, adicionalmente, aunque hay mucha rotación de personal en la compañía, el arquitecto es uno de los cargos que se conservan con más tiempo y por lo tanto da la opción de que se tenga mayor cantidad de información de los proyectos terminados y así poder brindar una información sustancial y valedera, ya que la primicia para un proyecto de minería de datos parte de la cantidad de datos y consistencia de información que se tenga; la idea finalmente es poder evaluar todos los proyectos que hayan sido culminados y de los cuales se pueda obtener la información requerida por medio de encuestas.

Se realizaron 83 encuestas para identificar el comportamiento que han tenido los 30 factores seleccionados en la ejecución de los proyectos.

➤ Tipo: Encuesta de pregunta cerrada usando escala de *Likert* de 1 a 5 para medir el comportamiento de los factores en los diversos proyectos.

➤ Locación: Encuesta online, usando la herramienta *Form* de la aplicación *Google Drive*. Dirección URL de la encuesta:

<https://docs.google.com/forms/d/1HWcLIm2mvxojfaHuU4uf1ECUO5vc0tbmEWdxFKZpkcA/viewform>

- Métodos de acceso: Internet, posterior a la ejecución de las encuestas, el resultado de estas se exportó a un archivo formato Excel xls, para darle manejo a la información.

Las preguntas realizadas en la encuesta contenía la información del factor que se iba a evaluar y la categoría al cual pertenece el factor, según el estudio bibliográfico realizado inicialmente, adicionalmente se adicionaron adjetivos a los factores con el fin de formular preguntas que facilitaran la respuesta por parte de los encuestados, quedando de la siguiente manera:

- Categoría: Capacidad de los desarrolladores, hace referencia a las aptitudes técnicas de los desarrolladores en el uso de herramientas, plataformas y conocimiento de los procesos de desarrollo:

Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo.

Amplio nivel de experiencia de los desarrolladores con la plataforma / entorno del proyecto.

Amplio nivel de experiencia de los desarrolladores con herramientas y lenguaje de desarrollo.

Amplios conocimientos de los desarrolladores con las herramientas usadas para soportar actividades de desarrollo del proyecto.

- Categoría: Conocimiento del dominio, hace referencia a la comprensión que tienen los involucrados en el proyecto sobre el enfoque, aspectos y entorno del proyecto:

Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto.

Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto.

Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar.

Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto.

Amplios conocimientos de los desarrolladores sobre el producto a desarrollar.

- Categoría: Comunicación, hace referencia a la capacidad que tienen los involucrados en el proyecto para transferir la información entre ellos mismos:

Gran capacidad de comunicación de los desarrolladores.

Gran capacidad de comunicación de los diseñadores.

- Categoría: Composición del equipo, hace referencia a como está conformado el equipo de trabajo del proyecto:

Excelente conocimientos técnicos de los desarrolladores.

Excelente conocimientos técnicos de los diseñadores.

Nivel de experiencia en programación de los desarrolladores.

Nivel de experiencia del equipo en proyectos de desarrollo.

Uso de personas con postgrado en algún área de TI en el proyecto.

Tamaño del equipo del proyecto.

El factor capacidad del equipo se dividió en dos factores, debido al interés de poder evaluar por separado tanto a los desarrolladores como a los diseñadores:

Gran capacidad de los desarrolladores para resolver problemas.

Gran capacidad de los diseñadores para resolver problemas.

- Categoría: Complejidad del producto, hace referencia al grado de dificultad del proyecto para ser desarrollado:

Amplia complejidad de las interfaces externas del producto a desarrollar.

Cantidad de interfaces externas del producto (tanto de hardware como software).

Tamaño del producto a desarrollar.

- Categoría: Madurez de Gestión de Proyectos, hace referente al buen funcionamiento, entendimiento y aplicación de buenas prácticas de los procesos para realizar la gestión de proyectos:

Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto.

Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto.

El factor Adherencia a los planes en dos factores, debido al interés de poder evaluar por separado el cumplimiento del plan respecto al proyecto y respecto a los procesos de la empresa:

Excelente cumplimiento con el plan de trabajo respecto al proyecto.

Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa.

- Categoría: Colaboración, hace referencia a las relaciones y participaciones que tienen las personas involucradas en el proyecto:

Nivel de continuidad y estabilidad del equipo del proyecto.

Adicional a los factores y categorías anteriormente descritas se definieron 3 factores adicionales y se creó la categoría ejecución del proyecto, con el fin de determinar los factores que parten de la definición de proyecto exitoso (tiempo, costo y alcance):

- Categoría: Ejecución del proyecto, hace referencia a cómo se ejecuta el proyecto en términos de costo, tiempo y alcance; esta categoría fue creada solo con el

único fin de recolectar la información solicitada, ya que a partir de esta información se va a determinar si un proyecto es exitoso o no:

Ejecución del proyecto dentro del tiempo esperado

Ejecución del proyecto dentro del presupuesto esperado

Ejecución del proyecto dentro del alcance esperado

4.4.2. Descripción de datos.

La calificación de los factores en las encuestas estaba dada por una escala de *Likert* de 1 a 5 en la mayoría de las preguntas, donde:

1: Totalmente en desacuerdo

2: Parcialmente en desacuerdo

3: Ni de acuerdo ni en desacuerdo.

4: Parcialmente de acuerdo

5: Totalmente de acuerdo

Tabla 1. Operacionalización de variables

Variable		Tipo	Definición	Categoría	Escala
Dependiente	Independiente				
Proyecto Exitoso		Cualitativa – No métrica	Indica si el proyecto se ha ejecutado dentro del tiempo, presupuesto y alcance esperado.		Nominal - Dicotómica
	Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo	Cualitativa – No métrica	Indica si los desarrolladores tenían un amplio conocimiento de los procesos de desarrollo cuando se ejecutó el proyecto	Capacidad de los desarrolladores	Ordinal (Likert)
	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto.	Cualitativa – No métrica	Indica si los desarrolladores tenían una amplia experiencia con la plataforma del proyecto cuando este fue ejecutado.	Capacidad de los desarrolladores	Ordinal (Likert)
	Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto.	Cualitativa – No métrica	Indica si los desarrolladores tenían una amplia experiencia con las herramientas y lenguajes de desarrollo usados para la ejecución del proyecto.	Capacidad de los desarrolladores	Ordinal (Likert)
	Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto.	Cualitativa – No métrica	Indica qué tanto conocimiento tenían los desarrolladores sobre las herramientas que soportaban las actividades del proyecto	Capacidad de los desarrolladores	Ordinal (Likert)

		métrica	cuando este fue ejecutado.		
	Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto.	Cualitativa – No métrica	Indica qué tanta capacidad tenían los desarrolladores para entender la arquitectura del proyecto y resolver los riesgos que este presentaba cuando fue ejecutado.	Conocimiento del dominio	Ordinal (Likert)
	Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto.	Cualitativa – No métrica	Indica si los desarrolladores habían desarrollado aplicaciones similares a la del proyecto que se ejecutó.	Conocimiento del dominio	Ordinal (Likert)
	Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	Cualitativa – No métrica	Indica qué tanto conocimiento tenían los desarrolladores referentes a los aspectos del negocio del proyecto que se ejecutó, representa el factor “conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar”.	Conocimiento del dominio	Ordinal (Likert)
	Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto.	Cualitativa – No métrica	Indica qué tanto conocimiento tenían los desarrolladores referentes al dominio del proyecto que se ejecutó.	Conocimiento del dominio	Ordinal (Likert)
	Amplios conocimientos de los desarrolladores sobre el producto a desarrollar.	Cualitativa – No métrica	Indica qué tanto conocimiento tenían los desarrolladores sobre el proyecto que se ejecutó.	Conocimiento del dominio	Ordinal (Likert)

	Gran capacidad de comunicación de los desarrolladores.	Cualitativa – No métrica	Indica qué tanta capacidad tenían los desarrolladores para comunicarse con su equipo de trabajo y los clientes en el proyecto que se ejecutó.	Comunicación	Ordinal (Likert)
	Gran capacidad de comunicación de los diseñadores.	Cualitativa – No métrica	Indica qué tanta capacidad tenían los diseñadores para comunicarse con su equipo de trabajo y los clientes en el proyecto que se ejecutó.	Comunicación	Ordinal (Likert)
	Gran capacidad de los desarrolladores para resolver problemas.	Cualitativa – No métrica	Indica qué tanta capacidad tenían los desarrolladores para afrontar y resolver los problemas que se presentaron en la ejecución del proyecto.	Composición del equipo	Ordinal (Likert)
	Excelentes conocimientos técnicos de los desarrolladores.	Cualitativa – No métrica	Indica si los desarrolladores tenían amplios conocimientos técnicos cuando se ejecutó el proyecto.	Composición del equipo	Ordinal (Likert)
	Gran capacidad de los diseñadores para resolver problemas.	Cualitativa – No métrica	Indica qué tanta capacidad tenían los diseñadores para afrontar y resolver los problemas que se presentaron en la ejecución del proyecto.	Composición del equipo	Ordinal (Likert)
	Excelentes conocimientos técnicos de los diseñadores.	Cualitativa – No métrica	Indica si los diseñadores tenían amplios conocimientos técnicos cuando se ejecutó el proyecto.	Composición del equipo	Ordinal (Likert)

		métrica			
	Nivel de experiencia en programación de los desarrolladores	Cualitativa – No métrica	<p>Indica el nivel de experiencia que tenían desarrolladores cuando se ejecutó el proyecto, evaluado en años de experiencia:</p> <p><=1 año en promedio: Calificación 1</p> <p>>1<=2 años en promedio: Calificación 2</p> <p>>2<=3 años en promedio: Calificación 3</p> <p>>3<=4 años en promedio: Calificación 4</p> <p>Superior a 4 años en promedio: Calificación 5.</p>	Composición del equipo	Ordinal (Likert)
	Nivel de experiencia del equipo en proyectos de desarrollo.	Cualitativa – No métrica	<p>Indica el nivel de experiencia que tenía el equipo en proyectos de desarrollo cuando se ejecutó el proyecto,</p> <p>Evaluado en años de experiencia:</p> <p><= 1 año: Calificación 1</p> <p>> 1<= 2 años: Calificación 2</p> <p>> 2 <= 3 años: Calificación 3</p> <p>> 3<= 4 años: Calificación 4</p> <p>> 4 años: Calificación 5.</p>	Composición del equipo	Ordinal (Likert)
	Uso de personas con postgrado en el proyecto.	Cualitativa – No métrica	Indica la cantidad de personas que tenía postgrado en el equipo cuando se ejecutó el	Composición del equipo	Ordinal

		métrica	<p>proyecto, evaluado de la siguiente forma:</p> <p>0 personas: Calificación 1</p> <p>1 persona: Calificación 2</p> <p>2 personas: Calificación 3</p> <p>3 personas: Calificación 4</p> <p>Superior a 4 personas: Calificación 5.</p>		(Likert)
	Tamaño del equipo del proyecto.	Cualitativa – No métrica	<p>Indica la cantidad de personas que tenía el equipo del proyecto cuando este fue ejecutado; evaluado en número de personas por proyecto</p> <p>1-3 personas: Calificación 1</p> <p>4-6 personas: Calificación 2</p> <p>7-9 personas: Calificación 3</p> <p>10-12 personas: Calificación 4</p> <p>Superior a 13 personas: Calificación 5.</p>	Composición del equipo	Ordinal (Likert)
	Amplia complejidad de las interfaces externas del producto a desarrollar.	Cualitativa – No métrica	Indica si las interfaces externas del proyecto tenían una amplia complejidad cuando se ejecutó el proyecto.	Complejidad del producto	Ordinal (Likert)
	Cantidad de interfaces externas del producto (tanto de hardware como software).	Cualitativa – No métrica	Indica el número de interfaces externas que tenía el proyecto que se ejecutó, teniendo en cuenta tanto las interfaces de hardware como	Complejidad del producto	Ordinal

		métrica	de software, evaluado de la siguiente forma: 0 interfaces: Calificación 1 1 interface: Calificación 2 2 interfaces: Calificación 3 3 interfaces: Calificación 4 Superior a 4 interfaces: Calificación 5.		(Likert)
	Tamaño del producto a desarrollar.	Cualitativa – No métrica	Indica qué tan grande era el tamaño del proyecto que se desarrolló, medido en casos de uso: Inferior a 2 CU: Calificación 1 Entre 3 y 4 CU: Calificación 2 Entre 5 y 6 CU: Calificación 3 Entre 7 y 8 CU: Calificación 4 Superior a 8 CU: Calificación 5.	Complejidad del producto	Ordinal (Likert)
	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto.	Cualitativa – No métrica	Indica qué tan buena fue la definición de roles y responsabilidades para afrontar las actividades del proyecto que se ejecutó.	Madurez de Gestión de Proyectos	Ordinal (Likert)
	Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto.	Cualitativa – No métrica	Indica qué tan bien se definieron las responsabilidades y los niveles de autoridad del proyecto cuando este se ejecutó,	Madurez de Gestión de Proyectos	Ordinal (Likert)

		métrica			
	Excelente cumplimiento con el plan de trabajo respecto al proyecto.	Cualitativa – No métrica	Indica qué tan bien fue el cumplimiento del plan de trabajo con respecto a lo planteado en el proyecto que se ejecutó,	Madurez de Gestión de Proyectos	Ordinal (Likert)
	Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa.	Cualitativa – No métrica	Indica qué tan bien fue el cumplimiento del plan de trabajo con respecto a los procesos de trabajo que rigen en la empresa cuando se ejecutó el proyecto,	Madurez de Gestión de Proyectos	Ordinal (Likert)
	Nivel de continuidad y estabilidad del equipo del proyecto	Cualitativa – No métrica	Indica qué tanta continuidad a tenido el equipo realizando proyectos juntos, evaluado en cantidad de proyectos en los que ha participado el equipo: 1 proyecto: Calificación 1 2 proyectos: Calificación 2 3 proyectos: Calificación 3 4 proyectos: Calificación 4 Superior a 5 proyectos: Calificación 5	Colaboración	Ordinal (Likert)
	Ejecución del proyecto dentro del tiempo esperado	Cualitativa – No métrica	No representa un factor propio de la investigación bibliográfica realizada, pero se evaluó con el fin de saber si el proyecto ejecutado fue dentro del tiempo esperado.	Ejecución del proyecto	Ordinal (Likert)

	Ejecución del proyecto dentro del presupuesto esperado	Cualitativa – No métrica	No representa un factor propio de la investigación bibliográfica realizada, pero se evaluó con el fin de saber si el proyecto ejecutado fue dentro del presupuesto esperado.	Ejecución del proyecto	Ordinal (Likert)
	Ejecución del proyecto dentro del alcance esperado	Cualitativa – No métrica	No representa un factor propio de la investigación bibliográfica realizada, pero se evaluó con el fin de saber si el proyecto ejecutado fue dentro del alcance esperado.	Ejecución del proyecto	Ordinal (Likert)

Nota fuente: Elaboración propia.

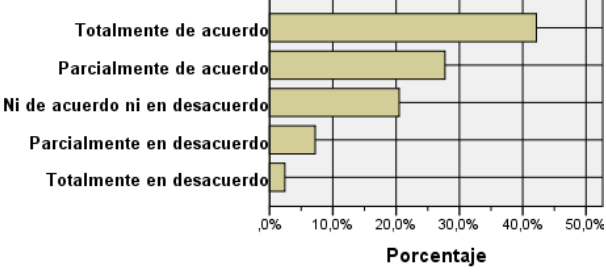
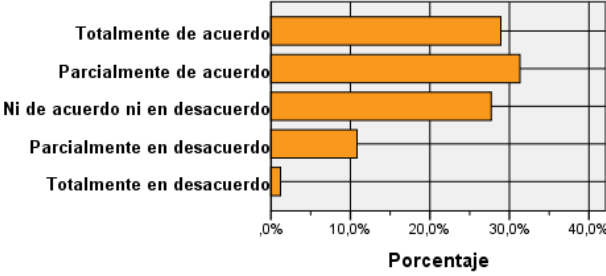
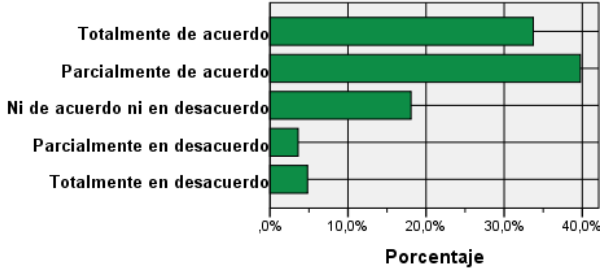
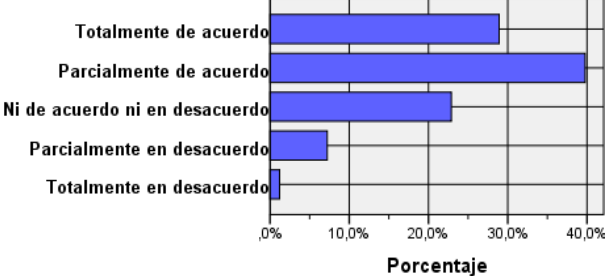
4.4.3. Exploración de datos.

A continuación se realiza una descripción y análisis de los datos encontrados en las encuestas:

- ✓ Factores que pertenecen a la categoría capacidad de los desarrolladores:

Capacidad de los desarrolladores hace referencia a las aptitudes técnicas de los desarrolladores en el uso de herramientas, plataformas y conocimiento de los procesos de desarrollo, en este sentido, en los 4 factores estudiados de esta categoría se presentó un predominio de buenas calificaciones donde la calificación predominante fue parcialmente de adecuado, a excepción del factor Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo el cual presento una calificación de totalmente de acuerdo en el 42% de los proyectos, así mismo se destaca que las calificaciones de parcialmente en desacuerdo y totalmente en desacuerdo juntas alcanzaron un porcentaje cercano al 10% en cada uno de los factores (ver Tabla 2).

Tabla 2. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría capacidad de los desarrolladores.

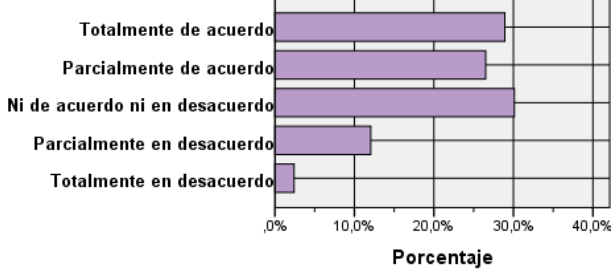
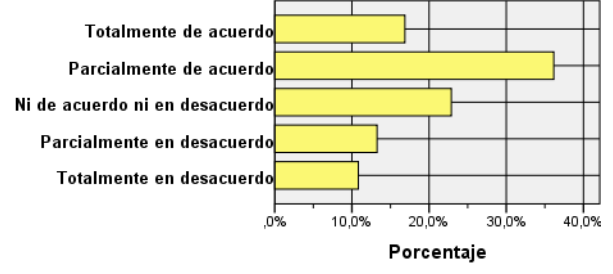
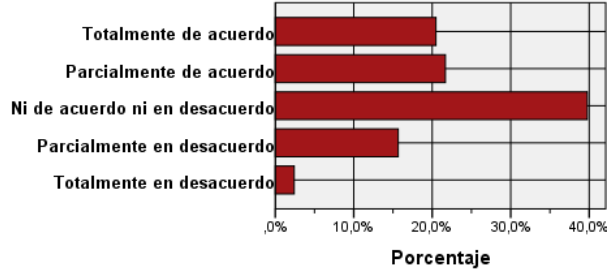
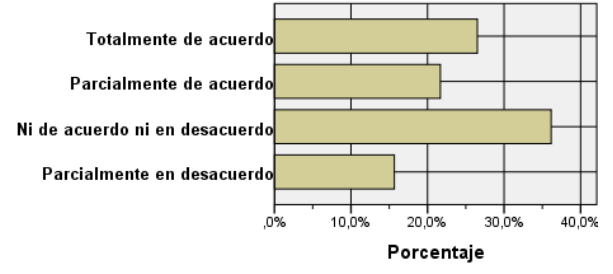
<p>Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>42,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>28,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>20,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>8,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	42,0%	Parcialmente de acuerdo	28,0%	Ni de acuerdo ni en desacuerdo	20,0%	Parcialmente en desacuerdo	8,0%	Totalmente en desacuerdo	2,0%
Categoría	Porcentaje												
Totalmente de acuerdo	42,0%												
Parcialmente de acuerdo	28,0%												
Ni de acuerdo ni en desacuerdo	20,0%												
Parcialmente en desacuerdo	8,0%												
Totalmente en desacuerdo	2,0%												
<p>Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>30,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>32,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>28,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>12,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	30,0%	Parcialmente de acuerdo	32,0%	Ni de acuerdo ni en desacuerdo	28,0%	Parcialmente en desacuerdo	12,0%	Totalmente en desacuerdo	2,0%
Categoría	Porcentaje												
Totalmente de acuerdo	30,0%												
Parcialmente de acuerdo	32,0%												
Ni de acuerdo ni en desacuerdo	28,0%												
Parcialmente en desacuerdo	12,0%												
Totalmente en desacuerdo	2,0%												
<p>Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>35,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>40,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>18,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>5,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>5,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	35,0%	Parcialmente de acuerdo	40,0%	Ni de acuerdo ni en desacuerdo	18,0%	Parcialmente en desacuerdo	5,0%	Totalmente en desacuerdo	5,0%
Categoría	Porcentaje												
Totalmente de acuerdo	35,0%												
Parcialmente de acuerdo	40,0%												
Ni de acuerdo ni en desacuerdo	18,0%												
Parcialmente en desacuerdo	5,0%												
Totalmente en desacuerdo	5,0%												
<p>Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>30,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>40,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>25,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>8,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	30,0%	Parcialmente de acuerdo	40,0%	Ni de acuerdo ni en desacuerdo	25,0%	Parcialmente en desacuerdo	8,0%	Totalmente en desacuerdo	2,0%
Categoría	Porcentaje												
Totalmente de acuerdo	30,0%												
Parcialmente de acuerdo	40,0%												
Ni de acuerdo ni en desacuerdo	25,0%												
Parcialmente en desacuerdo	8,0%												
Totalmente en desacuerdo	2,0%												

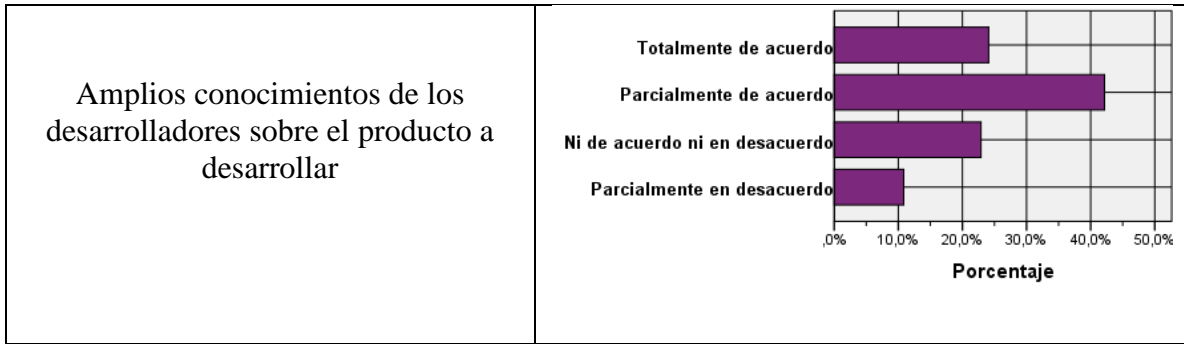
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

✓ Factores que pertenecen al conocimiento del dominio:

Conocimiento de dominio hace referencia a la comprensión que tienen los involucrados en el proyecto sobre el enfoque, aspectos y entorno del proyecto. En los 5 factores estudiados de esta categoría se presentó un predominio de calificaciones regulares (ni de acuerdo ni en desacuerdo) especialmente en las categorías Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar que alcanzo un total de 40% de los proyectos evaluados, Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto con 36% y Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto con un 30% . Las buenas calificaciones (parcialmente de acuerdo) fueron para los factores Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto que alcanzo un 36% de los proyectos estudiados y para el factor Amplios conocimientos de los desarrolladores sobre el producto a desarrollar que alcanzo un 42% de los proyectos; así mismo se destaca que los factores Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto y Amplios conocimientos de los desarrolladores sobre el producto a desarrollar no presentaron calificaciones de totalmente en desacuerdo (ver tabla 3).

Tabla 3. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría conocimiento del dominio.

<p>Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>30,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>28,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>30,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>12,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>3,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	30,0%	Parcialmente de acuerdo	28,0%	Ni de acuerdo ni en desacuerdo	30,0%	Parcialmente en desacuerdo	12,0%	Totalmente en desacuerdo	3,0%
Categoría	Porcentaje												
Totalmente de acuerdo	30,0%												
Parcialmente de acuerdo	28,0%												
Ni de acuerdo ni en desacuerdo	30,0%												
Parcialmente en desacuerdo	12,0%												
Totalmente en desacuerdo	3,0%												
<p>Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>18,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>35,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>23,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>13,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>11,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	18,0%	Parcialmente de acuerdo	35,0%	Ni de acuerdo ni en desacuerdo	23,0%	Parcialmente en desacuerdo	13,0%	Totalmente en desacuerdo	11,0%
Categoría	Porcentaje												
Totalmente de acuerdo	18,0%												
Parcialmente de acuerdo	35,0%												
Ni de acuerdo ni en desacuerdo	23,0%												
Parcialmente en desacuerdo	13,0%												
Totalmente en desacuerdo	11,0%												
<p>Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>22,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>23,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>40,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>16,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>3,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	22,0%	Parcialmente de acuerdo	23,0%	Ni de acuerdo ni en desacuerdo	40,0%	Parcialmente en desacuerdo	16,0%	Totalmente en desacuerdo	3,0%
Categoría	Porcentaje												
Totalmente de acuerdo	22,0%												
Parcialmente de acuerdo	23,0%												
Ni de acuerdo ni en desacuerdo	40,0%												
Parcialmente en desacuerdo	16,0%												
Totalmente en desacuerdo	3,0%												
<p>Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>27,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>22,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>36,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>15,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	27,0%	Parcialmente de acuerdo	22,0%	Ni de acuerdo ni en desacuerdo	36,0%	Parcialmente en desacuerdo	15,0%		
Categoría	Porcentaje												
Totalmente de acuerdo	27,0%												
Parcialmente de acuerdo	22,0%												
Ni de acuerdo ni en desacuerdo	36,0%												
Parcialmente en desacuerdo	15,0%												

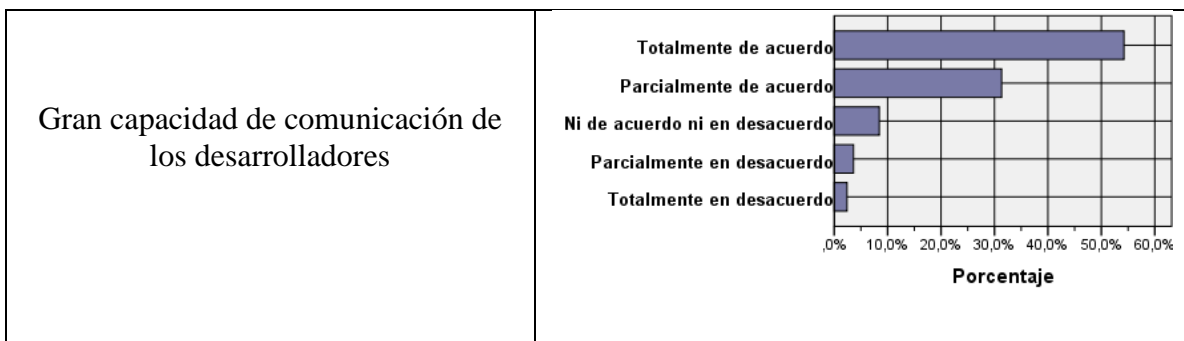


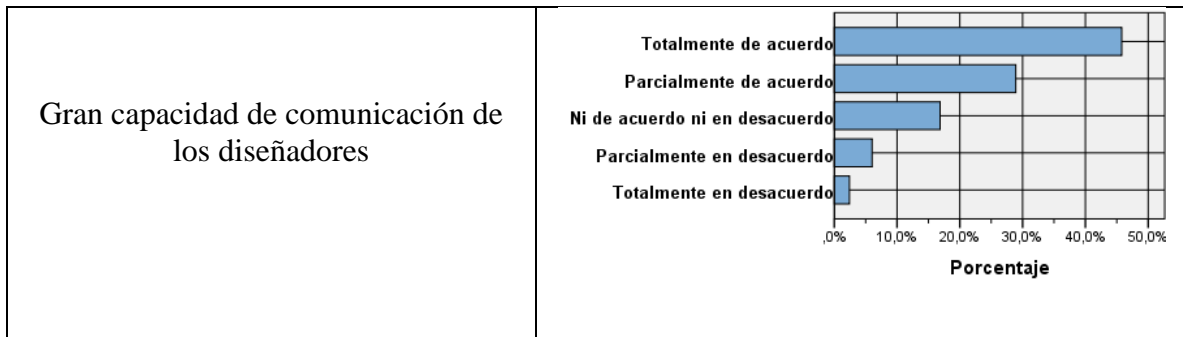
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

✓ Factores que representan la comunicación:

Estos factores hacen referencia a la capacidad que tienen los involucrados en el proyecto para transferir la información entre ellos mismos, estos dos factores que hace parte de esta categoría presentaron una distribución porcentual similar, ambas con excelentes calificaciones por encima de 45%, lo que indica que por lo menos el 45% de los proyectos fueron calificado de totalmente de acuerdo en ambos factores (ver tabla 4)

Tabla 4. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría comunicación



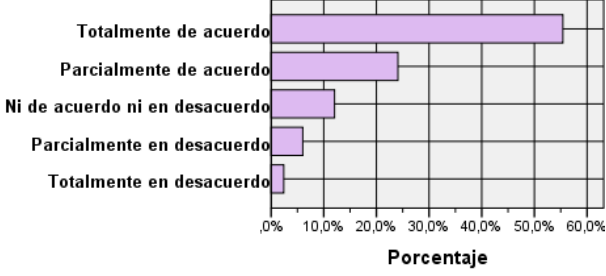
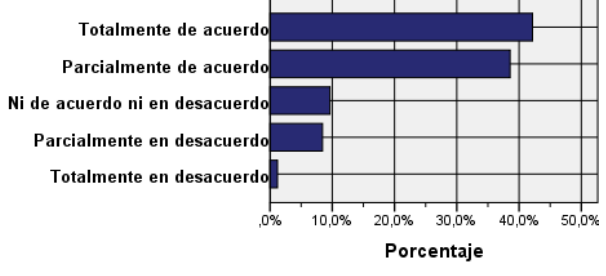
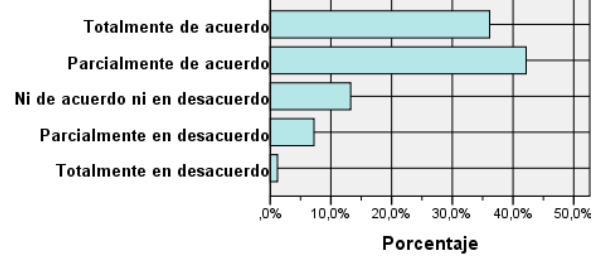
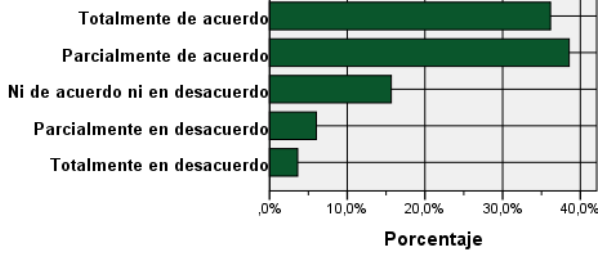


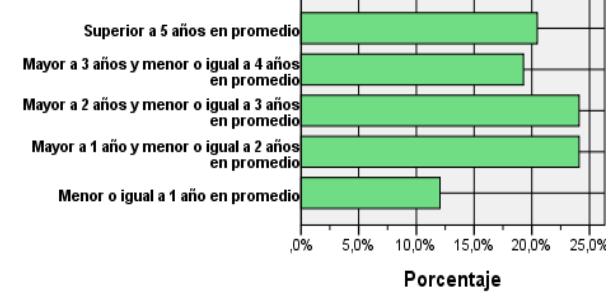
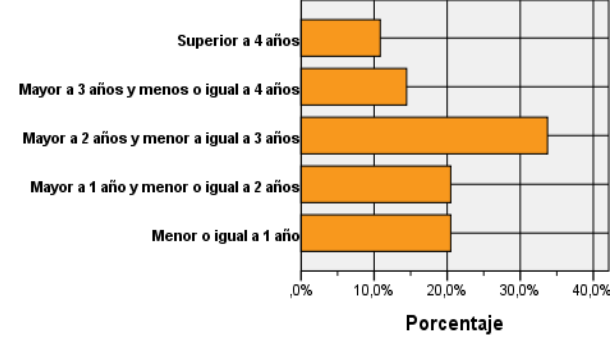
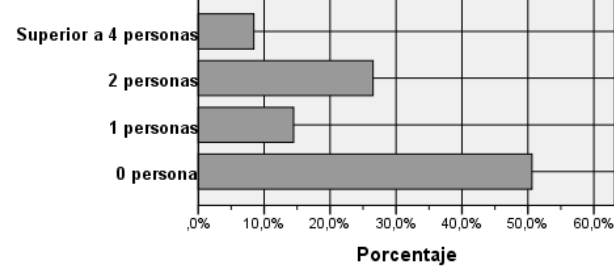
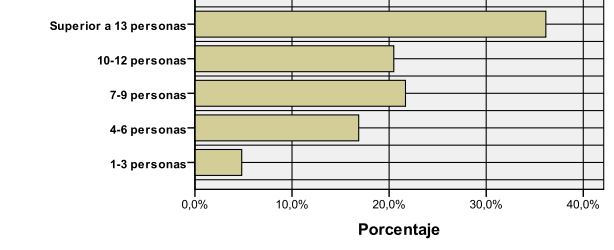
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

✓ Factores que representan la composición del equipo:

En la categoría hace referencia a como está conformado el equipo de trabajo del proyecto, en ella por lo menos un 75% de los proyectos evaluados en esta investigación fueron calificados de parcialmente de acuerdo o totalmente de acuerdo en los factores Gran capacidad de los desarrolladores para resolver problemas, Excelentes conocimientos técnicos de los desarrolladores, Gran capacidad de los diseñadores para resolver y Excelentes conocimientos técnicos de los diseñadores; En cuanto al Nivel de experiencia en programación de los desarrolladores se encontró que un 48% de los proyectos esta experiencia fue de 1 a 3 años, mientras 60% de los proyectos había alcanzado un Nivel de experiencia del equipo en proyectos de desarrollo de más de 2 años, Por su parte el tamaño del equipo de proyecto expresado en número de personas que trabajaron fue de 13 personas o más en el 36% de los casos, dentro de este personal por lo menos había un profesional con postgrado en alguna temática de TI en el 49% de los proyectos.(ver tabla 5).

Tabla 5. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría composición del equipo.

<p>Gran capacidad de los desarrolladores para resolver problemas</p>	 <table border="1"> <thead> <tr> <th>Calificación</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>55,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>25,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>10,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>5,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Calificación	Porcentaje	Totalmente de acuerdo	55,0%	Parcialmente de acuerdo	25,0%	Ni de acuerdo ni en desacuerdo	10,0%	Parcialmente en desacuerdo	5,0%	Totalmente en desacuerdo	2,0%
Calificación	Porcentaje												
Totalmente de acuerdo	55,0%												
Parcialmente de acuerdo	25,0%												
Ni de acuerdo ni en desacuerdo	10,0%												
Parcialmente en desacuerdo	5,0%												
Totalmente en desacuerdo	2,0%												
<p>Excelentes conocimientos técnicos de los desarrolladores</p>	 <table border="1"> <thead> <tr> <th>Calificación</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>45,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>40,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>10,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>8,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Calificación	Porcentaje	Totalmente de acuerdo	45,0%	Parcialmente de acuerdo	40,0%	Ni de acuerdo ni en desacuerdo	10,0%	Parcialmente en desacuerdo	8,0%	Totalmente en desacuerdo	2,0%
Calificación	Porcentaje												
Totalmente de acuerdo	45,0%												
Parcialmente de acuerdo	40,0%												
Ni de acuerdo ni en desacuerdo	10,0%												
Parcialmente en desacuerdo	8,0%												
Totalmente en desacuerdo	2,0%												
<p>Gran capacidad de los diseñadores para resolver</p>	 <table border="1"> <thead> <tr> <th>Calificación</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>38,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>42,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>12,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>6,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>2,0%</td> </tr> </tbody> </table>	Calificación	Porcentaje	Totalmente de acuerdo	38,0%	Parcialmente de acuerdo	42,0%	Ni de acuerdo ni en desacuerdo	12,0%	Parcialmente en desacuerdo	6,0%	Totalmente en desacuerdo	2,0%
Calificación	Porcentaje												
Totalmente de acuerdo	38,0%												
Parcialmente de acuerdo	42,0%												
Ni de acuerdo ni en desacuerdo	12,0%												
Parcialmente en desacuerdo	6,0%												
Totalmente en desacuerdo	2,0%												
<p>Excelentes conocimientos técnicos de los diseñadores</p>	 <table border="1"> <thead> <tr> <th>Calificación</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>35,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>38,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>15,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>6,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>3,0%</td> </tr> </tbody> </table>	Calificación	Porcentaje	Totalmente de acuerdo	35,0%	Parcialmente de acuerdo	38,0%	Ni de acuerdo ni en desacuerdo	15,0%	Parcialmente en desacuerdo	6,0%	Totalmente en desacuerdo	3,0%
Calificación	Porcentaje												
Totalmente de acuerdo	35,0%												
Parcialmente de acuerdo	38,0%												
Ni de acuerdo ni en desacuerdo	15,0%												
Parcialmente en desacuerdo	6,0%												
Totalmente en desacuerdo	3,0%												

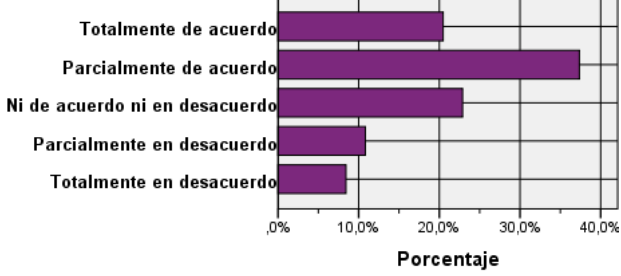
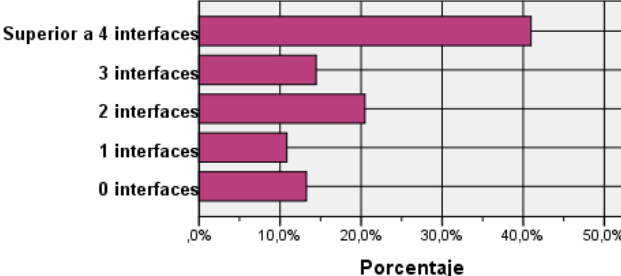
<p>Nivel de experiencia en programación de los desarrolladores</p>	 <table border="1"> <thead> <tr> <th>Nivel de experiencia</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Superior a 5 años en promedio</td> <td>~22%</td> </tr> <tr> <td>Mayor a 3 años y menor o igual a 4 años en promedio</td> <td>~20%</td> </tr> <tr> <td>Mayor a 2 años y menor o igual a 3 años en promedio</td> <td>~24%</td> </tr> <tr> <td>Mayor a 1 año y menor o igual a 2 años en promedio</td> <td>~24%</td> </tr> <tr> <td>Menor o igual a 1 año en promedio</td> <td>~12%</td> </tr> </tbody> </table>	Nivel de experiencia	Porcentaje	Superior a 5 años en promedio	~22%	Mayor a 3 años y menor o igual a 4 años en promedio	~20%	Mayor a 2 años y menor o igual a 3 años en promedio	~24%	Mayor a 1 año y menor o igual a 2 años en promedio	~24%	Menor o igual a 1 año en promedio	~12%
Nivel de experiencia	Porcentaje												
Superior a 5 años en promedio	~22%												
Mayor a 3 años y menor o igual a 4 años en promedio	~20%												
Mayor a 2 años y menor o igual a 3 años en promedio	~24%												
Mayor a 1 año y menor o igual a 2 años en promedio	~24%												
Menor o igual a 1 año en promedio	~12%												
<p>Nivel de experiencia del equipo en proyectos de desarrollo</p>	 <table border="1"> <thead> <tr> <th>Nivel de experiencia</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Superior a 4 años</td> <td>~10%</td> </tr> <tr> <td>Mayor a 3 años y menos o igual a 4 años</td> <td>~15%</td> </tr> <tr> <td>Mayor a 2 años y menor a igual a 3 años</td> <td>~35%</td> </tr> <tr> <td>Mayor a 1 año y menor o igual a 2 años</td> <td>~20%</td> </tr> <tr> <td>Menor o igual a 1 año</td> <td>~20%</td> </tr> </tbody> </table>	Nivel de experiencia	Porcentaje	Superior a 4 años	~10%	Mayor a 3 años y menos o igual a 4 años	~15%	Mayor a 2 años y menor a igual a 3 años	~35%	Mayor a 1 año y menor o igual a 2 años	~20%	Menor o igual a 1 año	~20%
Nivel de experiencia	Porcentaje												
Superior a 4 años	~10%												
Mayor a 3 años y menos o igual a 4 años	~15%												
Mayor a 2 años y menor a igual a 3 años	~35%												
Mayor a 1 año y menor o igual a 2 años	~20%												
Menor o igual a 1 año	~20%												
<p>Uso de personas con postgrado en el proyecto</p>	 <table border="1"> <thead> <tr> <th>Uso de personas con postgrado</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Superior a 4 personas</td> <td>~10%</td> </tr> <tr> <td>2 personas</td> <td>~28%</td> </tr> <tr> <td>1 personas</td> <td>~15%</td> </tr> <tr> <td>0 persona</td> <td>~52%</td> </tr> </tbody> </table>	Uso de personas con postgrado	Porcentaje	Superior a 4 personas	~10%	2 personas	~28%	1 personas	~15%	0 persona	~52%		
Uso de personas con postgrado	Porcentaje												
Superior a 4 personas	~10%												
2 personas	~28%												
1 personas	~15%												
0 persona	~52%												
<p>Tamaño del equipo del proyecto</p>	 <table border="1"> <thead> <tr> <th>Tamaño del equipo</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Superior a 13 personas</td> <td>~38%</td> </tr> <tr> <td>10-12 personas</td> <td>~22%</td> </tr> <tr> <td>7-9 personas</td> <td>~22%</td> </tr> <tr> <td>4-6 personas</td> <td>~18%</td> </tr> <tr> <td>1-3 personas</td> <td>~5%</td> </tr> </tbody> </table>	Tamaño del equipo	Porcentaje	Superior a 13 personas	~38%	10-12 personas	~22%	7-9 personas	~22%	4-6 personas	~18%	1-3 personas	~5%
Tamaño del equipo	Porcentaje												
Superior a 13 personas	~38%												
10-12 personas	~22%												
7-9 personas	~22%												
4-6 personas	~18%												
1-3 personas	~5%												

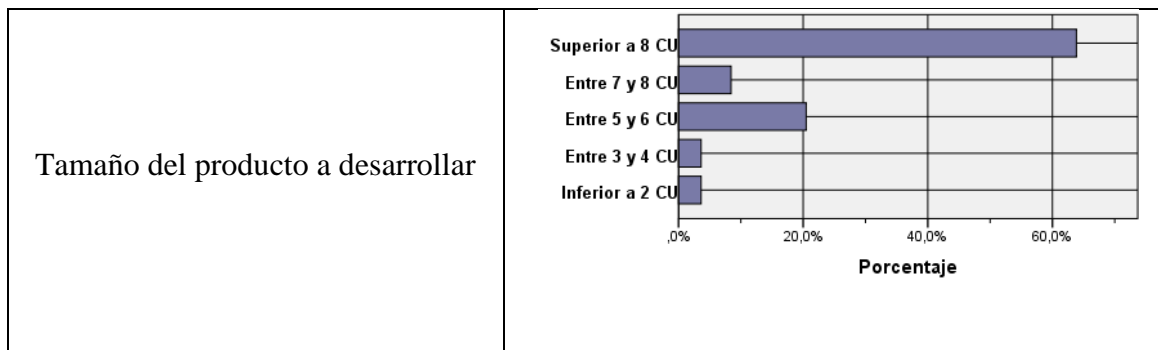
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

✓ Factores que representan la complejidad del producto:

Esta categoría hace referencia al grado de dificultad del proyecto para ser desarrollado. Y en ella 37% de los proyectos estudiados se encontraban parcialmente en desacuerdo con la amplia complejidad de las interfaces externas del producto a desarrollar, en el 62% de los casos el tamaño del proyecto a desarrollar presento más de 8 CU, mientras tanto, el 41% de los proyectos presento una cantidad de interfaces externas del producto (tanto de hardware como software) equivalentes a 4 interfaces (ver tabla 6).

Tabla 6. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría complejidad del producto.

<p>Amplia complejidad de las interfaces externas del producto a desarrollar</p>	 <table border="1"> <thead> <tr> <th>Calificación</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>21,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>37,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>23,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>11,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>8,0%</td> </tr> </tbody> </table>	Calificación	Porcentaje	Totalmente de acuerdo	21,0%	Parcialmente de acuerdo	37,0%	Ni de acuerdo ni en desacuerdo	23,0%	Parcialmente en desacuerdo	11,0%	Totalmente en desacuerdo	8,0%
Calificación	Porcentaje												
Totalmente de acuerdo	21,0%												
Parcialmente de acuerdo	37,0%												
Ni de acuerdo ni en desacuerdo	23,0%												
Parcialmente en desacuerdo	11,0%												
Totalmente en desacuerdo	8,0%												
<p>Cantidad de interfaces externas del producto (tanto de hardware como software)</p>	 <table border="1"> <thead> <tr> <th>Cantidad de interfaces</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Superior a 4 interfaces</td> <td>41,0%</td> </tr> <tr> <td>3 interfaces</td> <td>15,0%</td> </tr> <tr> <td>2 interfaces</td> <td>21,0%</td> </tr> <tr> <td>1 interfaces</td> <td>11,0%</td> </tr> <tr> <td>0 interfaces</td> <td>12,0%</td> </tr> </tbody> </table>	Cantidad de interfaces	Porcentaje	Superior a 4 interfaces	41,0%	3 interfaces	15,0%	2 interfaces	21,0%	1 interfaces	11,0%	0 interfaces	12,0%
Cantidad de interfaces	Porcentaje												
Superior a 4 interfaces	41,0%												
3 interfaces	15,0%												
2 interfaces	21,0%												
1 interfaces	11,0%												
0 interfaces	12,0%												

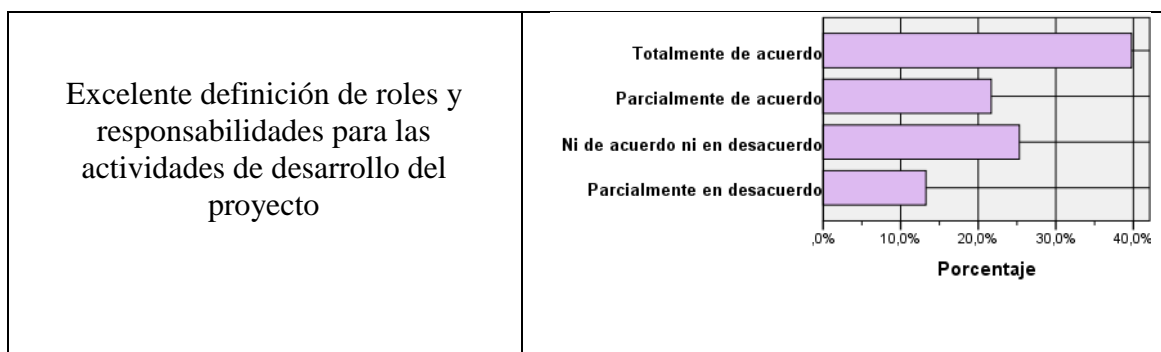


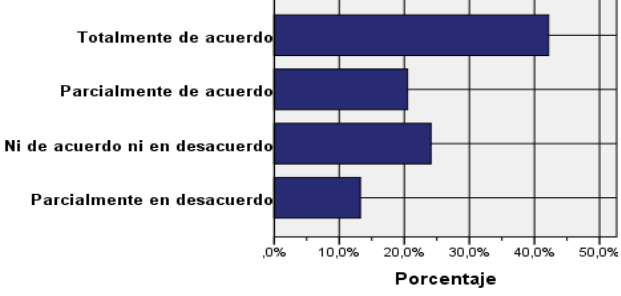
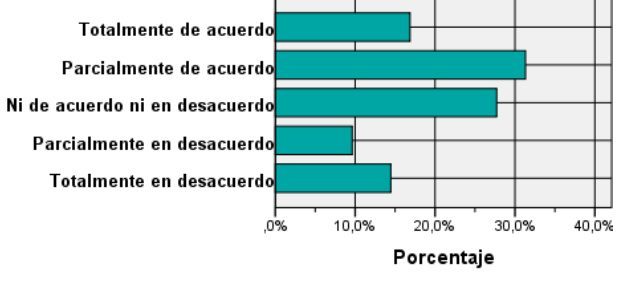
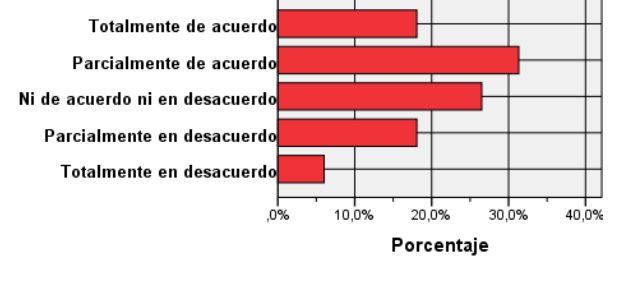
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

✓ Factores que representan la Madurez de gestión de proyectos:

Esta categoría hace referente al buen funcionamiento, entendimiento y aplicación de buenas prácticas de los procesos para realizar la gestión de proyectos. En esta categoría por lo menos el 40% de los proyectos presentaron una Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y una Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto; de igual forma, un porcentaje cercano al 30% de los proyectos presentaron Excelente cumplimiento con el plan de trabajo respecto al proyecto y Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa(Ver tabla 7).

Tabla 7. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría madurez de la gestión de proyectos.



<p>Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>45,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>20,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>25,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>10,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	45,0%	Parcialmente de acuerdo	20,0%	Ni de acuerdo ni en desacuerdo	25,0%	Parcialmente en desacuerdo	10,0%		
Categoría	Porcentaje												
Totalmente de acuerdo	45,0%												
Parcialmente de acuerdo	20,0%												
Ni de acuerdo ni en desacuerdo	25,0%												
Parcialmente en desacuerdo	10,0%												
<p>Excelente cumplimiento con el plan de trabajo respecto al proyecto</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>18,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>32,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>28,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>10,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>15,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	18,0%	Parcialmente de acuerdo	32,0%	Ni de acuerdo ni en desacuerdo	28,0%	Parcialmente en desacuerdo	10,0%	Totalmente en desacuerdo	15,0%
Categoría	Porcentaje												
Totalmente de acuerdo	18,0%												
Parcialmente de acuerdo	32,0%												
Ni de acuerdo ni en desacuerdo	28,0%												
Parcialmente en desacuerdo	10,0%												
Totalmente en desacuerdo	15,0%												
<p>Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>18,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>32,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>28,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>18,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>8,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	18,0%	Parcialmente de acuerdo	32,0%	Ni de acuerdo ni en desacuerdo	28,0%	Parcialmente en desacuerdo	18,0%	Totalmente en desacuerdo	8,0%
Categoría	Porcentaje												
Totalmente de acuerdo	18,0%												
Parcialmente de acuerdo	32,0%												
Ni de acuerdo ni en desacuerdo	28,0%												
Parcialmente en desacuerdo	18,0%												
Totalmente en desacuerdo	8,0%												

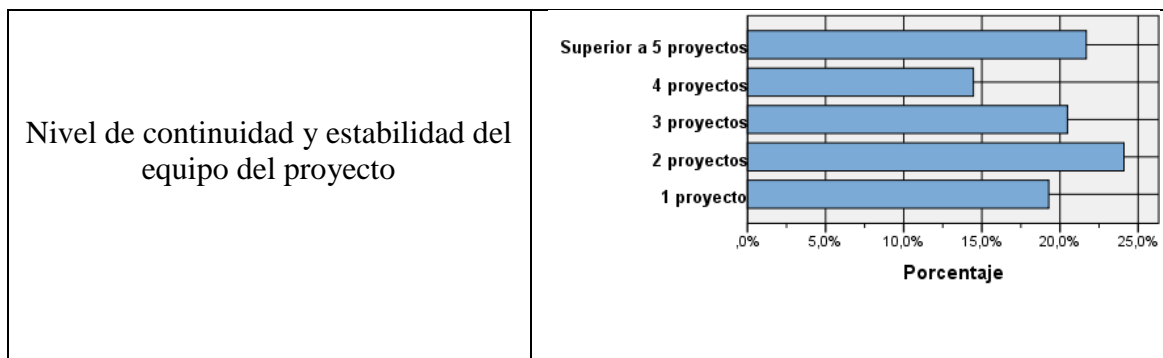
Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

✓ Factor que representa la colaboración del equipo:

Esta categoría hace referencia a las relaciones y participaciones que tienen las personas involucradas en el proyecto. El único factor que integro esta categoría fue Nivel de continuidad y estabilidad del equipo del proyecto y de él podemos decir que los proyectos evaluados se distribuyeron uniformemente dentro de los calificativos observados, de esta forma podemos decir que en el menor de los porcentajes (14.8%) de los proyectos tuvieron una continuidad y estabilidad de 4

proyectos. Es de especificar que este factor de esta categoría será incluido al momento de realizar un modelo multivariado (ver tabla 8).

Tabla 8. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría colaboración del equipo.

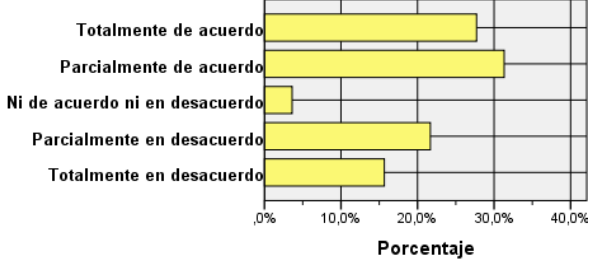
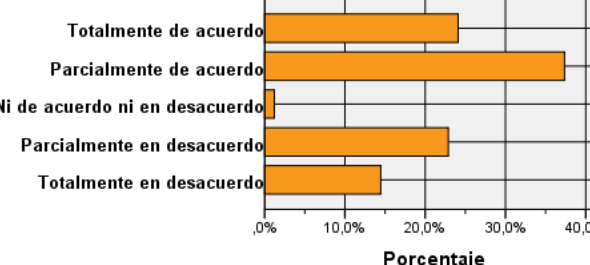
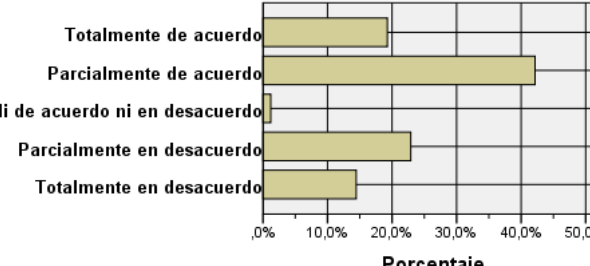


Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

- ✓ Factores que representan la ejecución del proyecto:

Esta categoría hace referencia a cómo se ejecuta el proyecto en términos de costo, tiempo y alcance. Al observar la tabla No 1. podemos establecer que se presentó un predominio de proyectos calificados como parcialmente de acuerdo en cuanto a la Ejecución del proyecto dentro del tiempo esperado, Ejecución del proyecto dentro del presupuesto esperado y Ejecución del proyecto dentro del alcance esperado alcanzado cada uno un 31%, 38% y 42% respectivamente. Es de recordar que la variable éxito del proyecto, que en esta investigación es la variable dependiente y cuyos valores definidos fueron éxito o fracaso, fue creada partir del promedio de estas tres factores, por lo tanto dichos factores no serán considerados en la construcción de un modelo multivariado posterior (ver tabla 9).

Tabla 9. Distribución porcentual de la calificación obtenida para cada factor dentro de la categoría ejecución del proyecto.

<p>Ejecución del proyecto dentro del tiempo esperado</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>28,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>31,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>5,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>22,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>15,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	28,0%	Parcialmente de acuerdo	31,0%	Ni de acuerdo ni en desacuerdo	5,0%	Parcialmente en desacuerdo	22,0%	Totalmente en desacuerdo	15,0%
Categoría	Porcentaje												
Totalmente de acuerdo	28,0%												
Parcialmente de acuerdo	31,0%												
Ni de acuerdo ni en desacuerdo	5,0%												
Parcialmente en desacuerdo	22,0%												
Totalmente en desacuerdo	15,0%												
<p>Ejecución del proyecto dentro del presupuesto esperado</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>25,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>38,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>2,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>23,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>15,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	25,0%	Parcialmente de acuerdo	38,0%	Ni de acuerdo ni en desacuerdo	2,0%	Parcialmente en desacuerdo	23,0%	Totalmente en desacuerdo	15,0%
Categoría	Porcentaje												
Totalmente de acuerdo	25,0%												
Parcialmente de acuerdo	38,0%												
Ni de acuerdo ni en desacuerdo	2,0%												
Parcialmente en desacuerdo	23,0%												
Totalmente en desacuerdo	15,0%												
<p>Ejecución del proyecto dentro del alcance esperado</p>	 <table border="1"> <thead> <tr> <th>Categoría</th> <th>Porcentaje</th> </tr> </thead> <tbody> <tr> <td>Totalmente de acuerdo</td> <td>18,0%</td> </tr> <tr> <td>Parcialmente de acuerdo</td> <td>42,0%</td> </tr> <tr> <td>Ni de acuerdo ni en desacuerdo</td> <td>2,0%</td> </tr> <tr> <td>Parcialmente en desacuerdo</td> <td>22,0%</td> </tr> <tr> <td>Totalmente en desacuerdo</td> <td>14,0%</td> </tr> </tbody> </table>	Categoría	Porcentaje	Totalmente de acuerdo	18,0%	Parcialmente de acuerdo	42,0%	Ni de acuerdo ni en desacuerdo	2,0%	Parcialmente en desacuerdo	22,0%	Totalmente en desacuerdo	14,0%
Categoría	Porcentaje												
Totalmente de acuerdo	18,0%												
Parcialmente de acuerdo	42,0%												
Ni de acuerdo ni en desacuerdo	2,0%												
Parcialmente en desacuerdo	22,0%												
Totalmente en desacuerdo	14,0%												

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.4.4. Verificar la calidad de los datos.

Con el fin de evitar el sesgo en las encuestas, la pérdida, errores y falta de datos, se hicieron las encuestas de manera *online* controlando la obligatoriedad de las preguntas en todo el cuestionario, así mismo, solo se tendrían datos en un rango de 1 a 5. De igual

forma se revisó que todos los datos resultados tuvieran valores comprendidos entre 1 y 5, y que no hubiera datos perdidos.

4.5. Preparación de los datos

4.5.1. Selección de datos y limpieza de datos.

Se excluyeron del análisis y modelamiento de los datos las variables conformadas por los factores: Ejecución del proyecto dentro del tiempo esperado, Ejecución del proyecto dentro del presupuesto esperado, Ejecución del proyecto dentro del alcance esperado, debido a que la información de estos factores se recolectó solo con el único fin de recrear la variable Proyecto Exitoso, adicionalmente no hacen parte de los factores que inciden en la inyección de defectos de desarrollo de software.

4.5.2. Construcción de datos.

Se construyó la variable Proyecto Exitoso, como el promedio de las variables Ejecución del proyecto dentro del tiempo esperado, ejecución del proyecto dentro del presupuesto esperado y ejecución del proyecto dentro del alcance esperado, se tomaron en cuenta solo

estas variables ya que la definición de éxito para esta investigación estaba contenida dentro del tiempo, costo y alcance.

Después de obtener el promedio, como el fin de la investigación es realizar una predicción de una variable dependiente partiendo de n variables independientes, donde se debe identificar si es exitoso o no, se determinó que si el promedio resultado era igual o superior a 3 el proyecto se calificaba con 1, que representaba exitoso, o si el resultado era inferior a 3 se calificaba con 0, que representaba fracaso.

4.5.3. Integración de datos.

Como solo se tiene una fuente de datos, que es el archivo Excel con el resultado de las encuestas, no se requiere hacer una combinación de datos.

4.6. Modelado

4.6.1. Seleccionar la técnica de modelado.

Se seleccionó el modelo de regresión logística binomial, se hizo teniendo en cuenta la naturaleza de la variable dependiente ya que es categórica binaria, es decir, que solo puede tomar dos valores, que para nuestro caso fue la variable Proyecto Exitoso cuyos

valores fueron (0) para proyectos no exitosos y (1) para los proyectos exitosos. Así mismo las variables independientes pueden ser de cualquier naturaleza ya sea cualitativa o cuantitativa.

Por otro lado la idea es construir un modelo y obtener una ecuación con fines de predicción, partiendo de la selección de n variables independientes.

4.6.2. Generar el diseño de las pruebas.

La prueba ómnibus valida si la varianza explicada en un conjunto de datos es significativamente mayor que la varianza no explicada, en general. Se asume que es estadísticamente significativa si el valor es ≤ 0.05 .

La prueba de bondad de ajuste va a indicar en el modelo si no existen diferencias significativas entre los valores observados y los pronosticados.

La prueba de Bartlett, la cual está referida a la matriz de correlaciones. Se contrasta la siguiente hipótesis nula (H_0): La matriz de correlaciones es una matriz de identidad; versus la hipótesis alternante: la matriz de correlaciones no es una matriz de identidad. En caso de rechazar la H_0 se concluye que las variables están correlacionadas entre sí, lo que da sentido al análisis componentes principales a realizar.

La prueba de Kaiser-Meyer-Olkin (KMO) estima un valor que de acuerdo a su ubicación en una escala permitirá concluir si el análisis realizado es conveniente. En la medida que los primeros sean más altos, el valor estimado estarán más cerca de uno, y por tanto el modelo factorial empleado será más efectivo.

Para interpretar la prueba de Bartlett se tiene en cuenta qué:

1 \geq KMO $>$ 0.9 excelente

0.9 \geq KMO $>$ 0.8 bueno

0.8 \geq KMO $>$ 0.7 aceptable

0.7 \geq KMO $>$ 0.6 regular

0.6 \geq KMO $>$ 0.5 deficiente

KMO \leq 0.5 inaceptable

Se utilizó el método de análisis de Spearman para el estudio de las correlaciones debido a la naturaleza de las variables involucradas, las cuales son de naturaleza cualitativa y cuyo nivel de medición es ordinal, y la idea es identificar los factores menos correlacionados entre sí. La correlación es una medida de asociación (fuerza) de la relación entre dos variables. Que varía de 0 (relación de azar) a 1 (relación lineal perfecta) o -1 (perfecta relación lineal negativa).

4.6.3. Construcción del modelo.

Para la construcción de los modelos de regresión logística se utilizó el 95% de los datos obtenidos lo que representó trabajar con 77 registros, los demás 6 registros serán utilizados para probar el modelo; los factores seleccionados para la construcción del modelo fueron: Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Tamaño del equipo del proyecto.

Ahora se presentan las funciones del análisis de regresión logística en la tabla 10. Se puede observar en ella que las variables incluidas en cada modelo son significativas (sig 0.000 y 0.004 inferior a 0.05); es decir, cada coeficiente (B) es diferente a cero, lo cual sugiere que dicha variable tiene un valor en cada función discriminante de regresión logística. Otro aspecto que corrobora lo anterior es el hecho de que el intervalo de confianza del 95% para $\text{Exp}(B)$, no es 1. Así mismo el signo de cada coeficiente (B) nos está sugiriendo la dirección en la cual dicha variable discrimina a los proyectos; por ejemplo en el modelo 2, el coeficiente (B) de 1.064 para la variable Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto nos indica que un aumento en una (1) unidad en esta variable, aumenta en un 1.064 la probabilidad de que un proyecto pertenezca al grupo de exitoso; y lo contrario sucede para el coeficiente (B) de -0.713 para la variable Tamaño del equipo del proyecto, la cual sugiere que un aumento en una (1) unidad en esta variable, disminuye en un 0.713 la probabilidad de que un proyecto pertenezca a este mismo grupo. Dicho en otras palabras, un proyecto exitoso tiende a ser diferenciado de uno no exitoso por tener de manera simultánea mayor Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Menor Tamaño del equipo del proyecto.

Tabla 10 Variables en la ecuación del modelo

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	.989	.260	14.468	1	.000	2.688	1.615	4.474
	Constante	-3.396	1.019	11.098	1	.001	.033		
Paso 2 ^b	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	1.064	.293	13.209	1	.000	2.898	1.633	5.143
	Tamaño del equipo del proyecto	-.713	.247	8.342	1	.004	.490	.302	.795
	Constante	-.997	1.306	.583	1	.445	.369		

a. Variable(s) introducida(s) en el paso 1: Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.6.3.1. Representación del modelo de regresión logística bivariado.

Partiendo de lo que se pretende mediante la Regresión Logística es expresar la probabilidad de que ocurra el evento en cuestión como función de ciertas variables, que se presumen relevantes o influyentes. Si ese hecho que queremos predecir lo representamos por Y (la variable dependiente), y las k variables explicativas (independientes y de control) se designan por X1, X2, X3,...,Xk, la ecuación general (o función logística) es:

1

$$P(Y=1) = \frac{\exp(-\alpha - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \dots - \beta_k X_k)}{1 + \exp(-\alpha - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \dots - \beta_k X_k)}$$

$$1 + \exp(-\alpha - \beta_1 X_1 - \beta_2 X_2 - \beta_3 X_3 - \dots - \beta_k X_k)$$

Donde $\alpha, \beta_1, \beta_2, \beta_3, \dots, \beta_k$ son los parámetros del modelo, y \exp denota la función exponencial. Esta función exponencial es una expresión simplificada que corresponde a elevar el número e a la potencia contenida dentro del paréntesis, siendo e el número o constante de Euler, o base de los logaritmos neperianos (cuyo valor aproximado a la milésima es 2,718).

Ahora bien, para nuestro modelo la función logística sería de la siguiente manera:

$$P(Y=1) = \frac{1}{1 + 2.718^{(-(-0.997) - (1.064 (X1)) - (-0.713(X2)))}}$$

Donde X1= Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto

X2 = Tamaño del equipo del proyecto

4.6.4. Evaluación del modelo.

La técnica utilizada para construir el modelo fue de paso a paso hacia adelante por ser el de mayor precisión, mediante el procedimiento de máxima verosimilitud (stepwise forward LR).

Para evaluar de forma global la validez modelo tenemos los indicadores de la tabla Resumen del modelo: el cual sugiere que en el 1 primer paso tenemos la inclusión

del primer factor (Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto) y el índice de R cuadrado de Nagelkerke, sugiere que dicha variable explica el 27.8% de la variación de la variable dependiente (calificación del proyecto), mientras en el segundo paso con la inclusión del segundo factor (Tamaño del equipo del proyecto) el modelo mejora ostensiblemente al alcanzar un 40.7% de la variación de la variable dependiente la cual es explicada por los factores incluidos en el modelo.

Tabla 11 Resumen del modelo de la regresión logística bivariado

Resumen del modelo			
Paso	-2 log de la verosimilitud	R cuadrado de Cox y Snell	R cuadrado de Nagelkerke
1	86,049 ^a	,206	,278
2	76,197 ^b	,301	,407

a. La estimación ha finalizado en el número de iteración 4 porque las estimaciones de los parámetros han cambiado en menos de ,001.

b. La estimación ha finalizado en el número de iteración 5 porque las estimaciones de los parámetros han cambiado en menos de ,001.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Los resultados de esta prueba arrojaron dos modelos de inclusión de variables: en el primer modelo se incluyó únicamente la variable Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto; en el segundo modelo se incluyó la variable, Tamaño del equipo del proyecto. Es de recordar que cada vez que se incluye una variable nueva en cada modelo, las demás permanecen en él; es decir en el

modelo dos quedaron las variables Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Tamaño del equipo del proyecto. Es importante señalar, que aunque el factor Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto presentó fue seleccionada para ser parte de este análisis, aquí en el análisis multivariado se excluyó por no aportar suficiente información para la construcción de los modelos, puntuación de 6,369 para una sig. 0.012

Tabla 12 Variables que no están en la ecuación

		Puntuación	gl	Sig.
Paso 0	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	16,930	1	,000
	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	6,369	1	,012
	Tamaño del equipo del proyecto	10,016	1	,002
	Estadísticos globales	25,372	3	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

En la Tabla 13 sobre los coeficientes del modelo podemos señalar que los dos modelos son estadísticamente significativos al tener un valor ≤ 0.02 ; es decir, las variables independientes tomadas juntas tienen un efecto discriminante altamente significativo sobre la variable dependiente. Es decir ayudan a predecir cual pertenece a éxito y cual a fracaso.

Tabla 13 Pruebas ómnibus sobre los coeficientes del modelo

	Chi cuadrado	gl	Sig.
Paso	17,755	1	,000
Paso 1 Bloque	17,755	1	,000
Modelo	17,755	1	,000
Paso	9,852	1	,002
Paso 2 Bloque	27,607	2	,000
Modelo	27,607	2	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

De igual forma la prueba de bondad de ajuste (ver tabla 14) sugiere que en cada modelo no existen diferencias significativas entre los valores observados y los pronosticados; lo que implica que en cada modelo estimado los datos observados y los datos esperados son similares en buena medida. Es de recordar, en este aspecto, que en la medida en que el valor de la significancia se encuentra mucho más alto de 0.05 (se acepta la hipótesis nula, el modelo propuesto se ajusta a lo observado), el ajuste de los datos es mucho mejor; y en este caso el primer modelo fue el que más se ajustó a los datos.

Tabla 14 Prueba de Hosmer y Lemeshow

Paso	Chi cuadrado	gl	Sig.
1	,516	2	,773
2	5,679	7	,578

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Para la evaluación del modelo se seleccionó al azar el 5% de los datos (6 proyectos) dado que para la construcción del modelo se utilizó el 95% de los datos restantes (77 proyectos), estas cantidades fueron seleccionadas con el fin de tener un grupo de proyectos que permitieran evaluar el modelo. Aplicando la ecuación que describe el

modelo, se obtuvo un valor que va de 0 a 1, para determinar si un proyecto es exitoso, este valor como mínimo deberá superar un valor de 0.8, mientras un proyecto no exitoso estará por debajo de dicha cifra, entre más cercano a 1 sea el resultado de aplicar el modelo más probabilidad de que el proyecto sea exitoso, por el contrario entre más cercano a 0 más probabilidad de que el proyecto no sea exitoso. Los resultados de realizar lo anterior se muestran en la tabla de contingencia 1 en ella se puede apreciar que el modelo pronostico acertó en la clasificación de proyecto exitoso en 1 de 2 dos, mientras en los proyectos no exitosos logro clasificar 3 de 4, en general el modelo logro clasificar correctamente 4 proyectos de 6, es decir, tuvo un 66.6% de aciertos.

Tabla de contingencia 1

Recuento		Calificación Proyecto		Total
		Fracaso	Exitoso	
Pronostico	Fracaso	3	1	4
	Exitoso	1	1	2
Total		4	2	6

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.6.4.1. Simulación.

Se realizó una simulación del modelo aplicándolo a toda la base de datos (resultados de las encuestas), se encontró que el modelo pronostico logro clasificar correctamente 37 proyectos exitosos de un total de 48, es decir, logro identificar correctamente el 77% de los proyectos exitosos, por otra parte, logro identificar 25/35, es decir, un 71% de los proyectos no exitosos, en general, el modelo pudo identificar correctamente 62

proyectos de 83 posibles, en otras palabras el modelo pudo identificar correctamente el 74% de los proyectos evaluados.

Simulación

Recuento		Calificación Proyecto		Total
		Fracaso	Exitoso	
		Pronostico	Fracaso	
	Exitoso	10	37	47
Total		35	48	83

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

4.7. Evaluación

4.7.1. Capacidad de los desarrolladores.

4.7.1.1. *Análisis correlacional entre los factores que pertenecen a la capacidad de los desarrolladores.*

Esta técnica de análisis es útil para estudiar las asociaciones de un conjunto de variables de escala u ordinales.

Basándonos en los niveles críticos (marcados en azul, ver tabla 14) podemos afirmar, en términos estadísticos, que en la presente investigación existe una múltiple correlación positiva entre todas las variables que forman la categoría capacidad de los desarrolladores

(sig. =0.000 para todas las relaciones), es de resaltar que la correlación con mayor fuerza la ofrecen los factores de Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto y Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto, al presentar un coeficiente de correlación(rho) de 0.786, lo cual indica que el 62% (rho²) de la variabilidad del factor Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto es explicado por el factor Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto.

Tabla 15 Matriz de correlaciones entre los factores que componen la categoría capacidad de los desarrolladores.

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto	Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto
	Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo	,740**	,715**	,605**
		.000	.000	.000
	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto		,786**	,719**
			.000	.000
	Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto			,604**
				.000

**La correlación es significativa al nivel 0,01 (bilateral).

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.1.2. Análisis de componentes principales de los factores que pertenecen a la capacidad de los desarrolladores.

Realizada la prueba de Bartlett, y por tener una Significancia estadística = 0.00, inferior al valor del alfa de 0.05 fijado apriori, se rechaza la Ho. Se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,841
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	208,537
	gl	6
	Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

Por lo tanto, la prueba de Bartlett calculada es igual a 0.841 alcanzando un nivel de bueno, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores intercorrelacionados significativamente.

En el cuadro de comunalidades (ver tabla 16), se miden el porcentaje de la varianza de cada una de los factores que se explica por el resto de las factores que componen la categoría, en el caso específico de esta investigación en la categoría capacidad de los desarrolladores, el factor Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto es explicado en un 82% por los demás 3 factores, lo cual reafirma los hallazgos del análisis correlacional. En síntesis, las comunalidades en

esta tabla son todas de valores altos, por lo tanto los 4 factores de esta categoría se encuentran ampliamente relaciones entre sí, lo que indica que el componente extraído representa bien los demás factores como se explicará más adelante.

Tabla 16 Cuadro de comunalidades capacidad de los desarrolladores

	Inicial	Extracción
Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo	1,000	,780
Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	1,000	,820
Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto	1,000	,789
Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto	1,000	,736
Método de extracción: Análisis de Componentes principales.		

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada (ver tabla 17) muestra un modelo que es explicado por un solo componente. En la columna " Sumas de las saturaciones al cuadrado de la extracción " se indica que el primer componente explica el 78.119% de la variación total de la categoría capacidad de los desarrolladores, lo que puede interpretarse como un porcentaje bueno, teniendo más de las $\frac{3}{4}$ partes de la variación de la capacidad de los desarrolladores explicada por los cuatro factores, es de anotar, que en este análisis solo se toman valores de la columna "Autovalores iniciales" cuya valor total para cada componente sea mayor de 1, así cada componente muestra los valores propios, que son la proporción de la varianza total en todas las variables que se explica por ese factor.

Tabla 17 Varianza total explicada capacidad de los desarrolladores

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,125	78,119	78,119	3,125	78,119	78,119
2	,371	9,269	87,388			
3	,269	6,732	94,121			
4	,235	5,879	100,000			

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

A continuación la Matriz de componentes, da el factor de cargas. Esta es la salida central para el análisis de componentes principales. El factor de cargas, también llamados saturaciones en componentes en análisis de componentes principales, son los coeficientes de correlación entre las variables y en este sentido el factor de mayor carga fue Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto que presentó un valor de 0.905, valor muy alto considerado, generalmente, de "excelente " de esta manera este es el factor de mayor importancia en la dicha categoría, por lo tanto, sería el factor candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente
	1
Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo	,883
Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	,905
Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto	,888
Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto	,858
Método de extracción: Análisis de componentes principales.	
a. 1 componentes extraídos	

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

4.7.2. Conocimiento del dominio.

4.7.2.1. *Análisis Correlacional entre los factores que pertenecen al conocimiento del dominio.*

Basándonos en los niveles críticos (marcados en azul, ver tabla 16) de la siguiente tabla podemos afirmar, en términos estadísticos, que en la presente investigación existe una múltiple correlación positiva entre todas las variables que forman la categoría conocimiento del dominio (sig. =0.000 para todas las relaciones), es de resaltar que la correlación con mayor fuerza la ofrecen los factores de Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar y Amplios

conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto, al presentar un coeficiente de correlación(rho) de 0.787, lo cual indica que el 61% (rho²) de la variabilidad del factor Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar es explicado por el factor Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto.

Tabla 18 Matriz de correlaciones entre los factores que componen la categoría conocimiento del dominio

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto	Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto	Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto	Amplios conocimientos de los desarrolladores sobre el producto a desarrollar
Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto		,561**	,525**	,389**	,507**
		.000	.000	.000	.000
Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto			,543**	,366**	,445**
			.000	.001	.000
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar				,787**	,705**
				.000	.000
Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto					,641**
					.000

**La correlación es significativa al nivel 0,01 (bilateral).

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.2.2. Análisis de componentes principales de los factores que pertenecen al conocimiento del dominio.

Realizada la prueba de Bartlett, se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.	,786
Chi-cuadrado aproximado	214,754
Prueba de esfericidad de Bartlett	
GI	10
Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.786 alcanzando un nivel de aceptable, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores intercorrelacionados significativamente.

En el cuadro de comunalidades (ver tabla 19), se miden el porcentaje de la varianza de cada una de los factores que se explica por el resto de las factores que componen la categoría, en el caso específico de esta investigación en la categoría conocimiento del dominio, el factor Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar es explicado en un 80% por los demás 4 factores, lo

cual reafirma los hallazgos del análisis correlacional. En síntesis, las comunalidades en esta tabla casi todos son valores medianos, es decir, el valor promedio esta en 0,55 contando de 0 a 1, lo que indica que el componente extraído representan de forma aceptable los demás factores, por lo tanto los 4 factores de esta categoría se encuentran relaciones entre si y todos ellos hacen parte de un solo componente como se explicará más adelante.

Tabla 19 Cuadro de comunalidades conocimiento del dominio

	Inicial	Extracción
Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto	1,000	,568
Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto	1,000	,515
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	1,000	,805
Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto	1,000	,651
Amplios conocimientos de los desarrolladores sobre el producto a desarrollar	1,000	,708

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada (ver tabla 20) muestra un modelo que es explicado por un solo componente. En la columna " Sumas de las saturaciones al cuadrado de la extracción " se indica que el primer componente explica el 64.95% de la variación total

de la categoría conocimiento del dominio, lo que puede interpretarse como un porcentaje bueno, teniendo un poco menos de las $\frac{3}{4}$ partes de la variación de la categoría conocimiento del dominio, explicada por los 5 factores que la componen, es de anotar, que en este análisis solo se toman valores de la columna “Autovalores iniciales” cuya valor total para cada componente sea mayor de 1, así cada componente muestra los valores propios, que son la proporción de la varianza total en todas las variables que se explica por ese factor.

Tabla 20 Varianza total explicada conocimiento del dominio

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,248	64,950	64,950	3,248	64,950	64,950
2	,817	16,345	81,295			
3	,427	8,536	89,832			
4	,328	6,550	96,382			
5	,181	3,618	100,000			

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial.

A continuación la matriz de componentes, da el factor de cargas. Esta es la salida central para el análisis de componentes principales. El factor de mayor carga fue Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar que presento un valor de 0.897, valor muy alto considerado, generalmente, de "excelente " de esta manera este es el factor de mayor importancia en la dicha categoría, por lo tanto, sería el factor candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente
	1
Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto	,754
Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto	,718
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	,897
Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto	,807
Amplios conocimientos de los desarrolladores sobre el producto a desarrollar	,842
Método de extracción: Análisis de componentes principales.	
a. 1 componentes extraídos	

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.3. Comunicación.

4.7.3.1. Análisis correlacional entre los factores que pertenecen a la comunicación.

De acuerdo con el nivel crítico (marcados en azul, ver tabla 16) de la tabla siguiente podemos afirmar, en términos estadísticos, que en la presente investigación existe una correlación positiva entre las factores que forman la categoría de comunicación (sig. =0.000), se presentó un coeficiente de correlación (rho) de 0.468, lo cual indica que el 21% (rho²) de la variabilidad del factor Gran capacidad de comunicación de los desarrolladores es explicado por la Gran capacidad de comunicación de los diseñadores.

Tabla 21 Matriz de correlaciones entre los factores que componen la categoría comunicación

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Gran capacidad de comunicación de los desarrolladores	Gran capacidad de comunicación de los diseñadores
Gran capacidad de comunicación de los desarrolladores		,468**
		.000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.3.2. *Análisis de componentes principales de los factores que pertenecen a la comunicación.*

Realizada la prueba de Bartlett, se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,500
Chi-cuadrado aproximado		27,238
Prueba de esfericidad de Bartlett	gl	1
	Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.500 alcanzando un nivel de inaceptable, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores muy poco

intercorrelacionados, esto es debido a que solo se presentaron dos factores y este tipo de análisis es ideal cuando se tienen más factores.

En el cuadro de comunalidades (ver tabla 22), se muestra lo identificado en el análisis correlacional en sentido de que el factor Gran capacidad de comunicación de los desarrolladores es explicado en un 76% por el factor Gran capacidad de comunicación de los diseñadores.

Tabla 22 Cuadro de comunalidades comunicación

	Inicial	Extracción
Gran capacidad de comunicación de los desarrolladores	1,000	,768
Gran capacidad de comunicación de los diseñadores	1,000	,768

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada (ver tabla 23) muestra un modelo que es explicado por un solo componente. En la columna " Sumas de las saturaciones al cuadrado de la extracción " se indica que el primer componente explica el 76.78% de la variación total de la categoría comunicación, lo que puede interpretarse como un porcentaje bueno, teniendo que las $\frac{3}{4}$ partes de la variación de la categoría comunicación, es explicada por los 2 factores que la componen, es de anotar, que en este análisis solo se tomas valores de la columna "Autovalores iniciales" cuya valor total para cada componente sea mayor de 1, así cada componente muestra los valores propios, que son la proporción de la varianza total en todas las variables que se explica por ese factor.

Tabla 23 Varianza total explicada comunicación

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	1,536	76,789	76,789	1,536	76,789	76,789
2	,464	23,211	100,000			
Método de extracción: Análisis de Componentes principales.						

A continuación la matriz de componentes, da el factor de cargas. Nos ofrece un coeficiente de 0.876 para cada uno de los factores, por lo tanto, cualquiera de los factores es candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente
	1
Gran capacidad de comunicación de los desarrolladores	,876
Gran capacidad de comunicación de los diseñadores	,876
Método de extracción: Análisis de componentes principales.	
a. 1 componentes extraídos	

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.4. Composición del equipo.

4.7.4.1. Análisis correlacional entre los factores que representan la composición del equipo.

Teniendo en cuenta los niveles críticos (marcados en azul, ver tabla 17) de la tabla siguiente podemos afirmar, en términos estadísticos, que en la presente investigación existe una múltiple correlación positiva entre todas las variables que forman la categoría composición del equipo (significancia que oscila entre 0.000 y 0.01), a excepción de los factores Uso de personas con postgrado en el proyecto y Tamaño del equipo del proyecto los cuales no presentaron tantas correlaciones con las demás factores.

Es de resaltar que la correlación con mayor fuerza la ofrecen los factores de Gran capacidad de los diseñadores para resolver problemas y Excelentes conocimientos técnicos de los diseñadores, al presentar un coeficiente de correlación (ρ) de 0.909, lo cual indica que el 82% (ρ^2) de la variabilidad del factor Gran capacidad de los diseñadores para resolver problemas es explicado por el factor Excelentes conocimientos técnicos de los diseñadores.

Tabla 24 Matriz de correlaciones entre los factores que componen la categoría composición del equipo

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Gran capacidad de los desarrolladores para resolver problemas	Excelentes conocimientos técnicos de los desarrolladores	Gran capacidad de los diseñadores para resolver problemas	Excelentes conocimientos técnicos de los diseñadores	Nivel de experiencia en programación de los desarrolladores	Nivel de experiencia del equipo en proyectos de desarrollo	Uso de personas con postgrado en el proyecto	Tamaño del equipo del proyecto
Gran capacidad de los desarrolladores para resolver problemas	,680**	,537**	,481**	,457**	,367**	.181	-.134	
	.000	.000	.000	.000	.001	.102	.226	
Excelentes conocimientos técnicos de los desarrolladores		,521**	,587**	,531**	,364**	,300**	-.094	
		.000	.000	.000	.001	.006	.400	
Gran capacidad de los diseñadores para resolver problemas			,909**	,438**	,253*	.190	-.114	
			.000	.000	.021	.085	.303	
Excelentes conocimientos técnicos de los diseñadores				,558**	,380**	,280*	-.062	
				.000	.000	.010	.580	
Nivel de experiencia en programación de los desarrolladores					,569**	,406**	.109	
					.000	.000	.326	
Nivel de experiencia del equipo en proyectos de desarrollo						,394**	,315**	
						.000	.004	
Uso de personas con postgrado en el proyecto							,245*	
							.026	
La correlación es significativa al nivel 0,01 (bilateral).								
La correlación es significativa al nivel 0,05 (bilateral).								

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.4.2. Análisis de componentes principales de los factores que representan la composición del equipo.

Realizada la prueba de Bartlett, se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.	,746
Chi-cuadrado aproximado	367,646
Prueba de esfericidad de Bartlett	
gl	28
Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.746 alcanzando un nivel de aceptable, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores intercorrelacionados significativamente.

En el cuadro de comunalidades (ver tabla 25), se miden el porcentaje de la varianza de cada una de los factores que se explica por el resto de las factores que componen la categoría, en el caso específico de esta investigación en la categoría composición del equipo, el factor Excelentes conocimientos técnicos de los diseñadores es explicado en un 80% por los demás factores, lo cual reafirma los hallazgos del análisis correlacional. En síntesis, las comunalidades en esta tabla presenta valores dispersos entre 0.504 y 0.800, lo que indica que son varios los componentes extraídos se explicará más adelante.

Tabla 25 Cuadro de comunalidades composición del equipo

	Inicial	Extracción
Gran capacidad de los desarrolladores para resolver problemas	1,000	,683
Excelentes conocimientos técnicos de los desarrolladores	1,000	,737
Gran capacidad de los diseñadores para resolver problemas	1,000	,775
Excelentes conocimientos técnicos de los diseñadores	1,000	,800
Nivel de experiencia en programación de los desarrolladores	1,000	,631
Nivel de experiencia del equipo en proyectos de desarrollo	1,000	,640
Uso de personas con postgrado en el proyecto	1,000	,504
Tamaño del equipo del proyecto	1,000	,622
Método de extracción: Análisis de Componentes principales.		

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada (ver tabla 26) muestra un modelo que es explicado por dos componentes. En la columna " Sumas de las saturaciones al cuadrado de la rotación " indica que el primer componente explica por sí solo el 43.760% de la variación total de la categoría composición del equipo, mientras el segundo componente solo aporta el 23.628% de la variación total de esta misma, en resumen ambos componentes juntos explican el 67.388% de la variación total de la categoría en cuestión.

Tabla 26 Varianza total explicada composición del equipo

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,895	48,683	48,683	3,895	48,683	48,683	3,501	43,760	43,760
2	1,496	18,705	67,388	1,496	18,705	67,388	1,890	23,628	67,388
3	,726	9,076	76,464						
4	,653	8,168	84,632						
5	,552	6,900	91,533						
6	,395	4,933	96,465						
7	,210	2,619	99,084						
8	,073	,916	100,000						

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

En la Matriz de componentes y Matriz de componentes de componentes rotados, se destaca que el factor Excelentes conocimientos técnicos de los diseñadores presento en ambas tablas el valor más alto, sin embargo para este análisis nos centraremos en la Matriz de componentes de componentes rotados en ella podemos apreciar varias cosas, entre ellas que en el componente 1 los factores que la integran son: Gran capacidad de los desarrolladores para resolver problemas, Excelentes conocimientos técnicos de los desarrolladores, Gran capacidad de los diseñadores para resolver problemas y Excelentes conocimientos técnicos de los diseñadores, explícitamente por presentar valores altos en dicho componente en la tabla matriz de componentes rotado; el factor Nivel de experiencia en programación de los desarrolladores y Nivel de experiencia del equipo en proyectos de desarrollo no serán tenidos en cuenta para ninguno de los dos componentes por tener valores similares en ambos componentes, finalmente, el factor Tamaño del equipo del proyecto con un valor de 0.757 es el más importante del componente dos. Por lo tanto, los factores Excelentes conocimientos técnicos de los diseñadores del componente 1 y Tamaño del equipo del proyecto del componente 2 serán los candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente	
	1	2
Gran capacidad de los desarrolladores para resolver problemas	,790	-,241
Excelentes conocimientos técnicos de los desarrolladores	,849	-,126
Gran capacidad de los diseñadores para resolver problemas	,796	-,376
Excelentes conocimientos técnicos de los diseñadores	,855	-,263
Nivel de experiencia en programación de los desarrolladores	,751	,259
Nivel de experiencia del equipo en proyectos de desarrollo	,653	,462
Uso de personas con postgrado en el proyecto	,428	,566
Tamaño del equipo del proyecto	,106	,781
Método de extracción: Análisis de componentes principales.		
a. 2 componentes extraídos		

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Matriz de componentes rotados

	Componente	
	1	2
Gran capacidad de los desarrolladores para resolver problemas	,820	,100
Excelentes conocimientos técnicos de los desarrolladores	,827	,229
Gran capacidad de los diseñadores para resolver problemas	,880	-,021
Excelentes conocimientos técnicos de los diseñadores	,888	,106
Nivel de experiencia en programación de los desarrolladores	,582	,541
Nivel de experiencia del equipo en proyectos de desarrollo	,409	,687
Uso de personas con postgrado en el proyecto	,162	,691
Tamaño del equipo del proyecto	-,220	,757
Método de extracción: Análisis de componentes principales.		
Método de rotación: Normalización Varimax con Kaiser.		
a. La rotación ha convergido en 3 iteraciones.		

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.5. Complejidad del producto.

4.7.5.1. *Análisis correlacional entre los factores que representan la complejidad del producto.*

En concordancia con la tabla No 18, podemos afirmar, en términos estadísticos, que en la presente investigación existe una múltiple correlación positiva entre todas las variables que forman la categoría complejidad del producto (sig. ≤ 0.05 para todas las

relaciones), es de resaltar que la correlación con mayor fuerza la ofrecen los factores de Amplia complejidad de las interfaces externas del producto a desarrollar y Cantidad de interfaces externas del producto (tanto de hardware como software), al presentar un coeficiente de correlación(rho) de 0.419, lo cual indica que el 17.5% (rho²) de la variabilidad del factor Amplia complejidad de las interfaces externas del producto a desarrollar es explicado por el factor Cantidad de interfaces externas del producto (tanto de hardware como software).

Tabla 27 Matriz de correlaciones entre los factores que componen la categoría complejidad del producto.

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Amplia complejidad de las interfaces externas del producto a desarrollar	Cantidad de interfaces externas del producto (tanto de hardware como software)	Tamaño del producto a desarrollar
Amplia complejidad de las interfaces externas del producto a desarrollar		,419**	,216*
		.000	.050
Cantidad de interfaces externas del producto (tanto de hardware como software)			,412**
			.000

** . La correlación es significativa al nivel 0,01 (bilateral).

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.5.2. Análisis de componentes principales de los factores que representan la complejidad del producto.

Realizada la prueba de Bartlett, se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,566
	Chi-cuadrado aproximado	32,404
Prueba de esfericidad de Bartlett	gl	3
	Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.566 alcanzando un nivel de deficiente, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores intercorrelacionados con poca significancia estadística, en otras palabras las relaciones entre estas variables es débil.

En el cuadro de comunalidades (ver tabla 28), se observa que el factor Cantidad de interfaces externas del producto (tanto de hardware como software) es explicado en un 72% por los demás 2 factores.

Tabla 28 Cuadro de comunalidades complejidad del producto

	Inicia l	Extracción
Amplia complejidad de las interfaces externas del producto a desarrollar	1,000	,543
Cantidad de interfaces externas del producto (tanto de hardware como software)	1,000	,722
Tamaño del producto a desarrollar	1,000	,439
Método de extracción: Análisis de Componentes principales.		

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada muestra un modelo que es explicado por un solo 1 componente. En la columna " Sumas de las saturaciones al cuadrado de la rotación " indica que este componente explica por sí solo el 56.7% de la variación total de la categoría complejidad del producto.

Tabla 29 Varianza total explicada complejidad del producto

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	1.703	56.778	56.778	1.703	56.778	56.778
2	.817	27.234	84.012			
3	.480	15.988	100.000			

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

La matriz de componentes, da el factor de cargas. Nos ofrece un coeficiente de 0.85 para el factor Cantidad de interfaces externas del producto (tanto de hardware como software), por lo tanto, este factor es candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente
	1
Amplia complejidad de las interfaces externas del producto a desarrollar	,737
Cantidad de interfaces externas del producto (tanto de hardware como software)	,850
Tamaño del producto a desarrollar	,662

Método de extracción: Análisis de componentes principales.

a. 1 componentes extraídos

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.6. Madurez de gestión de proyectos.

4.7.6.1. Análisis correlacional entre los factores que representan la Madurez de gestión de proyectos.

Centrándonos en los niveles críticos (marcados en azul, ver tabla 19) de la tabla siguiente podemos afirmar, en términos estadísticos, que en la presente investigación existe una múltiple correlación positiva entre todas las variables que forman la categoría Madurez de gestión de proyectos (sig. =0.000 para todas las relaciones), es de resaltar

que la correlación con mayor fuerza, casi perfecta, la ofrecen los factores de Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto, al presentar un coeficiente de correlación(ρ) de 0.958, lo cual indica que el 91% (ρ^2) de la variabilidad del factor Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto es explicado por el factor Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto.

Tabla 30 Matriz de correlaciones entre los factores que componen la categoría Madurez de gestión de proyectos

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto	Excelente cumplimiento con el plan de trabajo respecto al proyecto	Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto		,958**	,557**	,566**
		.000	.000	.000
Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto			,535**	,571**
			.000	.000
Excelente cumplimiento con el plan de trabajo respecto al proyecto				,905**
				.000

La correlación es significativa al nivel 0,01 (bilateral).

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.6.2. *Análisis de componentes principales de los factores que representan la Madurez de gestión de proyectos.*

Realizada la prueba de Bartlett, se concluye que es poco probable que la matriz de correlación sea una matriz de identidad, y por tanto la correlación entre las variables es estadísticamente significativa.

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,612
	Chi-cuadrado aproximado	391,652
Prueba de esfericidad de Bartlett	gl	6
	Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.841 alcanzando un nivel de regular, de acuerdo a la escala presentada, por tanto el modelo factorial empleado nos indica que nos encontramos con un conjunto de factores intercorrelacionados significativamente.

En el cuadro de comunalidades (ver tabla 31), se miden el porcentaje de la varianza de cada una de los factores que se explica por el resto de las factores que componen la categoría, en el caso específico de esta investigación en la categoría Madurez de gestión de proyectos, el factor Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto es explicado en un 79% por los demás 3 factores, lo cual reafirma los hallazgos del análisis correlacional. En síntesis, las comunalidades en esta tabla son todas de valores altos, lo que indica que el componente extraído representan bien los demás factores, por lo tanto los 4 factores de esta categoría se

encuentran ampliamente relaciones entre si y todos ellos hacen parte de un solo componente como se explicara más adelante.

Tabla 31 Cuadro de comunalidades madurez de gestión de proyectos

	Inicial	Extracción
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	1,000	,798
Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto	1,000	,785
Excelente cumplimiento con el plan de trabajo respecto al proyecto	1,000	,733
Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa	1,000	,769

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro Varianza total explicada (ver tabla 32) muestra un modelo que es explicado por un solo componente. En la columna " Sumas de las saturaciones al cuadrado de la extracción " se indica que el primer componente explica el 77.1% de la variación total de la categoría Madurez de gestión de proyectos, lo que puede interpretarse como un porcentaje bueno, teniendo más de las $\frac{3}{4}$ partes de la variación de la Madurez de gestión de proyectos, explicada por los cuatro factores.

Tabla 32 Varianza total explicada madurez de gestión de proyectos

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3,084	77,105	77,105	3,084	77,105	77,105
2	,786	19,658	96,763			
3	,098	2,462	99,225			
4	,031	,775	100,000			

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

La Matriz de componentes presentó un valor de 0.893, para el factor Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto, valor muy alto considerado, generalmente, de "excelente " de esta manera este es el factor de mayor importancia en la dicha categoría, por lo tanto, sería el factor candidato para la construcción de un modelo multivariado posterior.

Matriz de componentes

	Componente
	1
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	,893
Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto	,886
Excelente cumplimiento con el plan de trabajo respecto al proyecto	,856
Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa	,877
Método de extracción: Análisis de componentes principales.	
a. 1 componentes extraídos	

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.7.7. Selección de factores.

A continuación se presentan los factores representativos de cada categoría que fueron seleccionados para el análisis de componentes principales, es de mencionar, que con este análisis se pretende analizar de manera simultánea el efecto de estos factores, que fueron seleccionados mediante el análisis de componentes principales realizado anteriormente a cada categoría estudiada en esta investigación. Así mismo la importancia de este análisis multivariado radica en un mejor entendimiento del fenómeno objeto de estudio obteniendo información que a través de los métodos estadísticos univariados y bivariados sería casi imposible de conseguir.

Factores seleccionados para el análisis multivariado:

- ✓ Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto (Capacidad de los desarrolladores)
- ✓ Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar (Conocimiento del dominio)
- ✓ Gran capacidad de comunicación de los desarrolladores (Comunicación)
- ✓ Excelentes conocimientos técnicos de los diseñadores (Composición del equipo)
- ✓ Tamaño del equipo del proyecto (Composición del equipo)
- ✓ Cantidad de interfaces externas del producto tanto de hardware como software (Complejidad del producto)
- ✓ Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto (Madurez de gestión de proyectos)
- ✓ Nivel de continuidad y estabilidad del equipo del proyecto (Colaboración del equipo)

Con el fin de evitar la multicolinealidad (Que no exista relación lineal perfecta entre sus variables), se hace deseable que el nivel de tolerancia sea igual a 1, y el valor de FIV también sea igual a 1 ó en su defecto el nivel de tolerancia sea superior a 0,4 y $FIV < 5$, para tener un grado de multicolinealidad muy bajo, en nuestro caso (ver tabla 33), se puede evidenciar que el grado de multicolinealidad es muy bajo de acuerdo a los valores descritos. Arias (2008).

Tabla 33 Estadísticos de colinealidad

Modelo		Estadísticos de colinealidad	
		Tolerancia	FIV
1	(Constante)		
	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	,639	1,564
	Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	,745	1,341
	Gran capacidad de comunicación de los desarrolladores	,692	1,446
	Excelentes conocimientos técnicos de los diseñadores	,406	2,463
	Tamaño del equipo del proyecto	,603	1,658
	Cantidad de interfaces externas del producto (tanto de hardware como software)	,619	1,615
	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	,442	2,260

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Para esta parte final del análisis se tendrán en cuenta aquellos factores donde los niveles críticos no se encuentre marcados en azul en la tabla No 20, es decir no presenten correlaciones estadísticamente significativas con otros factores, o en su defectos que dichas correlaciones no sean suficientemente fuertes (evitar multicolinealidad), en este sentido las factores que presentan estas características fueron están divididos en dos grupos: Tamaño del equipo del proyecto (Composición del equipo) y Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto (Madurez de gestión de proyectos), ó Cantidad de interfaces externas del producto tanto

de hardware como software (Complejidad del producto), y Excelentes conocimientos técnicos de los diseñadores (Composición del equipo), los posteriores análisis nos ayudaron a determinar cuál grupo de variables seleccionar.

Tabla 34 Matriz de correlaciones entre los factores seleccionados para la construcción de un modelo multivariado.

MATRIZ DE CORRELACIONES Rho de SPEARMAN	Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	Gran capacidad de comunicación de los desarrolladores	Excelentes conocimientos técnicos de los diseñadores	Tamaño del equipo del proyecto	Cantidad de interfaces externas del producto (tanto de hardware como software)	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	Nivel de continuidad y estabilidad del equipo del proyecto
	.000	.000	.000	.868	.493	.000	.000	
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar		.268*	.367**	-.150	.056	.383**	.209	
		.014	.001	.175	.616	.000	.058	
Gran capacidad de comunicación de los desarrolladores			.284**	-.104	.011	.372**	.267*	
			.009	.352	.921	.001	.015	
Excelentes conocimientos técnicos de los diseñadores				-.062	.007	.739**	.172	
				.580	.947	.000	.121	
Tamaño del equipo del proyecto					.589**	-.120	.093	
					.000	.282	.402	
Cantidad de interfaces externas del producto (tanto de hardware como software)						-.042	.347**	
						.705	.001	
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto							.199	
							.071	

La correlación es significativa al nivel 0,01 (bilateral).

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,697
	Chi-cuadrado aproximado	205,291
Prueba de esfericidad de Bartlett	gl	28
	Sig.	,000

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Por lo tanto, la prueba de Bartlett calculada es igual a 0.697 alcanzando un nivel de regular, de acuerdo a la escala presentada, por tanto el modelo (análisis) factorial presentado a continuación nos indica que nos encontramos que las relaciones existentes entre estos factores son consideradas débiles.

En el cuadro de comunalidades (ver tabla 34), se miden el porcentaje de la varianza de cada una de los factores que se explica por el resto de las factores analizados, en el caso específico de esta investigación el factor Cantidad de interfaces externas del producto (tanto de hardware como software) es explicado en un 80% por los demás factores.

Tabla 35 Cuadro de comunalidades factores análisis multivariado

	Inicial	Extracción
Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	1,000	,611
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	1,000	,412
Excelentes conocimientos técnicos de los diseñadores	1,000	,665
Tamaño del equipo del proyecto	1,000	,669
Cantidad de interfaces externas del producto (tanto de hardware como software)	1,000	,803
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	1,000	,651
Gran capacidad de comunicación de los desarrolladores	1,000	,488
Nivel de continuidad y estabilidad del equipo del proyecto	1,000	,463

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

El cuadro varianza total explicada (ver tabla 35) muestra un modelo que es explicado por dos componentes. En la columna " Sumas de las saturaciones al cuadrado de la rotación " indica que el primer componente explica por sí solo el 37.2% de la variación total del grupo de factores seleccionados, mientras el segundo componente solo aporta el 22.2% de la variación total de esta misma, en resumen ambos componentes juntos explican el 69.5% de la variación total.

Tabla 36 Varianza total explicada factores análisis multivariado

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	3.000	37.500	37.500	3.000	37.500	37.500	2.981	37.263	37.263
2	1.763	22.036	59.537	1.763	22.036	59.537	1.782	22.273	59.537
3	.911	11.385	70.922						
4	.714	8.926	79.848						
5	.571	7.136	86.983						
6	.482	6.023	93.006						
7	.310	3.881	96.887						
8	.249	3.113	100.000						

Método de extracción: Análisis de Componentes principales.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

En la Matriz de componentes y Matriz de componentes de componentes rotados, se destaca que el factor Cantidad de interfaces externas del producto (tanto de hardware como software) presento en ambas tablas el valor más alto, sin embargo para este análisis nos centraremos en la Matriz de componentes rotados en ella podemos apreciar varias cosas, entre ellas que en el componente 1 los factores que la integraron con mayor fuerza fueron: Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto, Excelentes conocimientos técnicos de los diseñadores y Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto; por su parte, los factores más importante del componente dos fueron Tamaño del equipo del proyecto y Cantidad de interfaces externas del producto (tanto de hardware como software). En esta ocasión los factores que presentaron mayores valores no fueron incluidos para la construcción del modelo multivariado, esto se debe a que esos factores están más ampliamente relacionados con los demás factores y se puede incurrir en

multicolinealidad, esto difiere a los análisis de episodios anteriores debido a que en esos momentos precisamente se necesitaba un factor que representara las categorías a las que pertenecía.

Matriz de componentes

	Componente	
	1	2
Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	,773	,115
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	,636	-,089
Gran capacidad de comunicación de los desarrolladores	,688	-,120
Excelentes conocimientos técnicos de los diseñadores	,807	-,117
Tamaño del equipo del proyecto	-,072	,815
Cantidad de interfaces externas del producto (tanto de hardware como software)	,112	,889
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	,790	-,163
Nivel de continuidad y estabilidad del equipo del proyecto	,480	,482

Método de extracción: Análisis de componentes principales.

a. 2 componentes extraídos.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Matriz de componentes rotados

	Componente	
	1	2
Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto	,753	,210
Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar	,642	-,009
Gran capacidad de comunicación de los desarrolladores	,698	-,034
Excelentes conocimientos técnicos de los diseñadores	,815	-,016
Tamaño del equipo del proyecto	-,172	,800
Cantidad de interfaces externas del producto (tanto de hardware como software)	,002	,896
Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	,804	-,064
Nivel de continuidad y estabilidad del equipo del proyecto	,417	,538

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

a. La rotación ha convergido en 3 iteraciones.

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

4.8. Despliegue

La decisión de usar el modelo presentado en el siguiente punto para ayudar a predecir el éxito o fracaso de un proyecto es decisión de la empresa, sin embargo, es conveniente incluir dentro de los procesos de la empresa la ejecución del modelo con el fin de que pueda dar una pequeña idea inicial de lo que se puede presentar en la ejecución de todo el proyecto, así mismo, se propone generar el modelo en un archivo Excel (como herramienta de sistematización) para facilitar la ejecución de este por parte de los gerentes de proyecto.

4.9. Análisis de resultados

Variables en la ecuación del modelo

		B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95% para EXP(B)	
								Inferior	Superior
Paso 1 ^a	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	.989	.260	14.468	1	.000	2.688	1.615	4.474
	Constante	-3.396	1.019	11.098	1	.001	.033		
Paso 2 ^b	Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto	1.064	.293	13.209	1	.000	2.898	1.633	5.143
	Tamaño del equipo del proyecto	-.713	.247	8.342	1	.004	.490	.302	.795
	Constante	-.997	1.306	.583	1	.445	.369		

a. Variable(s) introducida(s) en el paso 1: Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Finalmente, se puede decir que en este apartado que los dos modelos (paso 1 representa modelo 1 y paso 2 representa modelo 2) de la tabla variables en la ecuación del modelo sirven para discriminar o diferenciar proyectos exitosos de aquellos que no lo son, pero sólo se requiere uno y no más, y para ello la tabla de clasificación presenta la clasificación que realizó para cada modelo. En ella se observa que los 2 modelos tienen un porcentaje global de diferenciación de aproximadamente de 71%, pero varían al momento de clasificar correctamente a cada proyecto según su grupo. En el primer modelo, el porcentaje de diferenciación en cada grupo varía de 61.3% a 78.3% (Es de recordar que en este modelo la única variable incluida fue Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto). En cuanto al segundo modelo presentó un porcentaje global de discriminación de 74%, pero en dicho modelo,

se incluyeron las variables Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Tamaño del equipo del proyecto, este modelo puede diferenciar mejor a los proyectos exitosos (80%) que a aquellos proyectos no exitosos (64%). Por lo tanto se sugiere que el modelo 2, es el que mejor diferenciación realiza entre los grupos, dado que su porcentaje global de aciertos de 74% es mayor que el modelo 1; además con un modelo de diferenciación del 80.4% para proyectos éxitos es mayor que el del modelo 1.

Tabla 37 Tabla de clasificación

	Observado		Pronosticado		
			Calificación Proyecto		Porcentaje correcto
			Fracaso	Éxito	
Paso 1	Calificación Proyecto	Fracaso	19	12	61,3
		Éxito	10	36	78,3
	Porcentaje global				71,4
Paso 2	Calificación Proyecto	Fracaso	20	11	64,5
		Éxito	9	37	80,4
	Porcentaje global				74,0

a. El valor de corte es ,600

Nota fuente: Elaboración propia. Tomado en base al resultado obtenido en la herramienta SPSS trial

Para el segundo modelo de los 20+11 proyectos que fueron fracaso, 20 fueron pronosticados como fracaso, es decir, un porcentaje de acierto del 64.5%.

De los 9+37 proyectos exitosos, 37 fueron pronosticados como exitosos, lo que da un porcentaje de acierto del 80.4%

5. CONCLUSIONES

- La respuesta a la formulación del problema ¿Qué factores influyen en la consecución de proyectos exitosos en la empresa de *outsourcing* de TI de la ciudad de Medellín?, de acuerdo al estudio realizado, indica que los factores que inciden en la consecución de proyectos exitosos son: Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto, Tamaño del equipo de proyecto.

- A pesar de que los factores que se analizaron inciden en la inyección de defectos en el desarrollo de software según el estudio bibliográfico realizado, se puede determinar que los dos factores Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto y Tamaño del equipo de proyecto si se les da un buen manejo y aplican el modelo respectivo, ayudaría a reducir el número de inyección de defectos en el desarrollo de software, ya que por sí mismo ayudan a conseguir el éxito de los proyectos.
- La presente investigación proveerá información muy importante a los gerentes de proyecto de la compañía, que servirá de apoyo para la toma de decisiones y les ayudara a determinar qué acciones tomar antes de iniciar un proyecto con el fin de contribuir al éxito de este, Así mismo, proveerá información a la compañía en aras de contribuir al mejoramiento de los procesos
- El resultado de las encuestas permite concluir que todavía falta mucho trabajo por realizar en la empresa con el fin de aumentar las capacidades y conocimiento de los desarrolladores en los proyectos de desarrollo de software que permitan tener equipos de trabajo altamente competitivos, ya que se pudo observar que solo el factor capacidad de comunicación de los desarrolladores supero en un 50% el resultado de los encuestados donde respondieron que estaban totalmente de acuerdo, según la escala definida.

- La presente investigación presenta un beneficio muy grande para la empresa de TI de la ciudad de Medellín, debido a que, aunque diversos estudios realizados por los autores como Leyva, Rosado & Febles (2010), en una organización de desarrollo de software en Cuba; donde indican que los factores de éxito de un proyecto están determinados por la definición del alcance y prioridades del proyecto, la participación de usuario final, y el grado de compromiso de los clientes; o por González & Rodenes (2007), en la industria de software en México, que indicaron que los factores de éxito de un proyecto de software son Apoyo del Gobierno, Capital Humano, Calidad, Marketing e Innovación, o finalmente por Yousef, Gamal, Warda & M.Mahmoud (2006), en la industria de software egipcia, donde determinaron factores de éxito como requisitos, patrocinador del proyecto, y los clientes; dan prueba de la existencia de una gran variedad de factores de éxito, que si bien, pueden o no influir en el desarrollo de todos los proyectos de software a nivel mundial, o solo de forma directa en el mercado objeto de estudio, pero que difieren en gran medida a los identificados en la empresa de la presente investigación; al poder determinar cuáles aplican directamente a la compañía, da un valor agregado que permitirá fortalecer más los equipos de trabajo y las determinaciones que se deban tomar al momento de iniciar un nuevo proyecto y contribuir a su éxito.
- El análisis de componentes principales permitió en primera instancia reafirmar la correlación existente entre los factores que componen cada categoría y segunda instancia reducir los factores que la componen, identificando el factor más relevante que debería ser seleccionado a la hora de realizar un análisis

multivariado. Finalmente en el modelo multivariado permitió reducir el número de dimensiones necesarias que permitieran explicar el éxito o fracaso de un proyecto.

- El modelo de regresión logística generado permitió determinar que los factores que inciden en el éxito de un proyecto están determinados por:

Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto, Tamaño del equipo de proyecto; siendo estos los 2 únicos factores de los 27 estudiados capaces de explicar el éxito o fracaso de los proyectos.

- El modelo 2 presentado en el análisis de resultados finalmente se puede utilizar como base para ayudar a predecir si un proyecto es exitoso o no, pero no puede ser considerado como única fuente de consulta, ya que no tiene capacidad de predicción en un 100%, esto debido a las variables y a la cantidad de datos utilizados, ya que hay muchísimos más factores que intervienen en el desarrollo de un proyecto los cuales no fueron considerados en el presente estudio.

- La simulación realizada a la base de datos (resultado de las encuestas) permitió evaluar el modelo, logrando identificar correctamente el 77% de los proyectos exitosos y el 71% de los proyectos no exitosos.

- Finalmente se hizo una validación del modelo generado con el Msc Carlos Eduardo Ríos, quien tiene más de 10 años de experiencia trabajando con equipos

de desarrollo de software, considera que el resultado de los factores donde se tiene que: si se hace una excelente definición de roles y responsabilidades para las actividades de desarrollo y se trabaja con equipos de desarrollo pequeños, si es muy consistente en que se tenga una aproximación grande de tener éxito en el proyecto, ya que considera que si el equipo es pequeño el alcance es acorde a la capacidad del equipo, se puede realizar una mejor estimación de los artefactos, en los equipos pequeños hay una mayor cohesión y los factores de riesgo disminuyen, se tiene claridad de roles en las fases y se logra tener algún tipo de especialización por fases de estos roles y responsabilidades, en caso contrario a si se tiene un proyecto muy grande donde generalmente se requiere de un equipo igualmente grande y allí se presentan muchas dificultades en la estimación, es más difícil el control de los riesgos, la comunicación del equipo, se pueden transponer responsabilidades, se hace un poco más complicada la gestión y desarrollo del proyecto; considera que esto se lleva acabo si se tiene una muy buena metodología de desarrollo. Finalmente la mayoría de los proyectos exitosos donde él ha trabajado se ha tenido definición de roles y responsabilidades en las actividades del proyecto y fueron desarrollados por equipos de trabajo pequeños no superiores a 5 personas.

6. RECOMENDACIONES

Sería importante en un trabajo futuro lograr poder sistematizar la generación y análisis del modelo, permitiendo ingresar nuevas variables y datos con el fin de que se actualice frecuentemente y así poder ayudar a tomar mejores decisiones a los gerentes de la compañía.

Se recomienda para futuras investigaciones, aplicar la base de datos utilizada en el presente estudio a otros métodos de selección de variables y extracción de factores que permitan generar nuevos modelos y realizar comparaciones exhaustivas con el fin de identificar el mejor modelo que pueda dar una solución más eficiente a la necesidad del mercado de poder tener proyectos de software exitosos.

En caso de querer actualizar el modelo 2 se recomienda migrar la información a otra herramienta de uso libre, o en su defecto, realizar la compra del SPSS, ya que este es software propietario, y para realizar la investigación se utilizó la versión *trial* del mismo.

Sería importante que para un trabajo futuro se puedan analizar un mayor número de variables y datos, con el fin de que permitan generar un % mucho mayor de precisión en el modelo

El modelo 2 de la sección análisis de resultados se puede utilizar como base para ayudar a tomar decisiones en la compañía, pero no debe ser la única fuente de determinación para decidir si se desarrolla o no el proyecto, se deben considerar otras variables debido aunque a pesar de que se tuvo un % de precisión alto, tener un 18% de desconfianza también es un valor alto a considerar.

BIBLIOGRAFÍA

Arias, B. (2008). Desarrollo de un ejemplo de análisis factorial confirmatorio con Lisrel, AMOS y SAS. Seminario de Actualización en Investigación sobre Discapacidad SAID 2008. Recuperado de <http://www.benitoarias.com/articulos/afc.pdf>

Carroll, C. (7 de marzo de 2013). IT Success and Failure — the Standish Group CHAOS Report Success Factors. Blog Crhis F Carroll. Recuperado el 2 de Mayo de 2013, de <http://www.cafe-encounter.net/p1183/it-success-and-failure-the-chaos-report-factors>.

Clark, B. & Zubrow, D. (2001). How good is the software: a review of defect prediction techniques. Institutional repository - Carnegie Mellon University. Recuperado de <http://www.sei.cmu.edu/library/assets/defect-prediction-techniques.pdf>

Cobos, C., Zúñiga, J., Guarín, J., León, E. & Mendoza, M. (2010). CMIN - herramienta case basada en CRISP-DM para el soporte de proyectos de minería de datos. *Ingeniería e Investigación*, 30(3) 45-56. Recuperado de <http://www.redalyc.org/articulo.oa?id=64316140004>.

De la Fuente Fernández, S. (2011). Análisis factorial. Universidad Autónoma de Madrid. Recuperado de <http://www.fuenterrebollo.com/Economicas/ECONOMETRIA/MULTIVARIANTE/FACTORIAL/analisis-factorial.pdf>

Eveleens, J.L. & Verhoef, C. (2010). The Rise and Fall of the Chaos Report Figures. *IEEE Software*, vol. 27, no. 1, pp-30-36.

Gómez, V., & Abásolo, J.E (2008). Using Data Mining to Describe Long Hospital Stays. *Paradigma*, 3 (1). Recuperado de http://paradigma.uniandes.edu.co/index.php?option=com_content&view=article&catid=40%3Aarticulos&id=85%3Ausing-data-mining-to-describe-long-hospital-stays&lang=es

Fernández Valdés, M. M. & Punjuante Dante, G. (2008). Análisis conceptual de las principales interacciones entra la gestión de información, la gestión documental y la gestión del conocimiento. *ACIMED*, Vol. 18, n.1. Recuperado de http://scielo.sld.cu/scielo.php?pid=S1024-94352008000700007&script=sci_arttext&tlng=pt

González Bañales, D.L. & Rodenes Adam, M. (2007). Factores críticos de éxito de la industria del software y su relación con la orientación estratégica de negocio: un estudio empírico-exploratorio. *JISTEM: Journal of Information Systems and Technology Management*, Sin mes, 47-70.

González Cardona, J.C. (2011). *Sistema de apoyo para la acreditación de la calidad de programas académicos de la universidad de Caldas aplicando técnicas de minería de datos*. (Tesis de maestría) Repositorio institucional – Universidad Autónoma de Manizales. Recuperado de: <http://hdl.handle.net/11182/350>.

Gutiérrez, J.E. (2012). *Descubrimiento de conocimiento en la bases de datos académica de la universidad Autónoma de Manizales aplicando redes neuronales*. (Tesis de maestría). Repositorio institucional – Universidad Autónoma de Manizales. Recuperado de: [http://repositorio.autonoma.edu.co/jspui/bitstream/11182/345/1/MS-C-JGutierrez-Informe\(V2\).pdf](http://repositorio.autonoma.edu.co/jspui/bitstream/11182/345/1/MS-C-JGutierrez-Informe(V2).pdf)

Hernández Orallo, J. Ramírez Quintana, M.J & Ferri Ramírez, C. (2004). Introducción a la minería de datos. Madrid: Person Prentice Hall.

IBM Corporation. (2010). SPSS: Modeling: CRISP DM 1.0 - Step-by-step data mining guide. <http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=YTW03084USEN>.

Jacobs, J., Moll, J.V., Krause, P., Kusters, R., Trienekens, J. & Brombacher, A. (2005). Exploring defect causes in products developed by virtual teams, *Information and Software Technology*, Volume 47, Issue 6, Pages 399-410, Recuperado de <http://www.sciencedirect.com/science/article/pii/S0950584904001351>.

Jacobs, J., Moll, J.V. Kusters, R., Trienekens, J. & Brombacher, A. (2007). Identification of factors that influence defect injection and detection in development of software intensive products, *Information and Software Technology*, Volume 49, Issue 7, Pages 774-789. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0950584906001200>.

Leszak, M., Perry, D.E. & Stoll, D. (2002). Classification and evaluation of defects in a project retrospective, *Journal of Systems and Software*, Volume 61, Issue 3, Pages 173-187. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0164121201001467>.

Leyva Vázquez, M.Y., Rosado Rosello, R. & Febles Estrada, A. (2012). Modelado y análisis de los Factores Críticos de Éxito de los proyectos de software mediante Mapas Cognitivos Difusos. *Ciencias de la Información*, Mayo-Agosto, 41-46.

Linberg, K.R. (1999). Software developer perceptions about software project failure: a case study, *Journal of Systems and Software*, Volume 49, Issues 2–3, 30, Pages 177-192.

Recuperado de <http://www.sciencedirect.com/science/article/pii/S0164121299000941>

Mas Machuca, M. & Martinez Costa, C. (2008). Análisis del factor estratégico para alcanzar el éxito de un proyecto de gestión del conocimiento. Aplicación al sector de la consultoría.

Dirección y organización, DYO – 37. Recuperado de <http://www.revistadyo.es/index.php/dyo/article/view/38>

Merchan, L. (2007). Estudio de Factores Críticos de Éxito Local e Internacional para Empresas de la Industria del Software. *Avances*. Vol 4 N°3. Recuperado de <http://www.revistas.unal.edu.co/index.php/avances/article/viewFile/9935/10467>

Nizam Nasir, M.H. & Sahibuddin, S. (2011). Critical success factors for software projects: A comparative study. *Scientific Research and Essays*. Vol. 6(10), pp. 2174-2186.

Recuperado de <http://www.academicjournals.org/sre/pdf/pdf2011/18May/Nasir%20and%20Sahibuddin.pdf>

Pereira, J., Cerpa, N., & Rivas, M. (2004). Factores de éxito en proyectos de desarrollo de software: análisis de la industria chilena del software. In Workshop de Ingeniería de

Software, Arica. Recuperado de http://www.academia.edu/915974/Factores_de_exito_en_proyectos_de_desarrollo_de_software_analisis_de_la_industria_chilena_de_software.

Ramón, G. (sa). Correlación entre variables, Apuntes de clase del curso Seminario Investigativo VI. Repositorio institucional – Universidad de Antioquia. Recuperado de http://viref.udea.edu.co/contenido/menu_alterno/apuntes/ac34-marco-teorico-biblio.pdf

Reyes, F., Cerpa, N., Candia-Véjar, A. & Bardeen, M. (2011). The optimization of success probability for software projects using genetic algorithms, *Journal of Systems and Software*, Volume 84, Issue 5, Pages 775-785. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0164121210003456>

Reyes Matus, F. (2006). Una Herramienta para la Predicción de Riesgos en Proyectos de Software usando Modelos en Redes Bayesianas. (Tesis, Universidad De Talca). <http://campuscurico.otalca.cl/~fmeza/cursos/2007-1/pmt/memorias/FranciscoReyes.pdf>

Rout , T.P. & Abel , D.E. (1992). A model for defect insertion and detection in software development software, Presented at Proceedings of the 16th Australian Computer Science Conference, Brisbane, Australia: Griffith University. Recuperado de <http://www.sqi.griffith.edu.au/docs/sqi/misc/defmodel.pdf>

Soto Jaramillo, C.M. (2009). *Incorporación de técnicas multivariantes en un sistema gestor de base de datos*. (Tesis de maestría, Universidad Nacional de Colombia Sede Medellín). Recuperado de http://www.bdigital.unal.edu.co/895/1/71335481_2009.pdf

Standish Group. (2013). The Chaos Manifesto 2013. Recuperado el 20 de Septiembre de 2014, de <http://www.versionone.com/assets/img/files/CHAOSManifesto2013.pdf>.

Stutzke , M.A. & Smidts , C.S.. (2001). A stochastic model of fault introduction and removal during software development. *IEEE Software*, Volume 50, Issue 2, pp-184-193.

Recuperado

de

http://ieeexplore.ieee.org/xpl/articleDetails.jsp?tp=&arnumber=963126&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D963126

Terradez Gurra, M (s.a). Análisis de componentes principales, Proyecto e-Math. Repositorio institucional – Universidad Oberta de Catalunya. Recuperado de http://www.uoc.edu/in3/emath/docs/Componentes_principales.pdf

UNCTAD. (2012). Information economy report 2012, *The software industry and developing countries*. New York and Geneva. United Nations Publications. Recuperado de: http://unctad.org/en/PublicationsLibrary/ier2012_en.pdf.

Yousef, A. H., Gamal, A., Warda, A & Mahmoud, M. (2006). Projects Success Factors Identification using Data Mining . In *Computer Engineering and Systems, The 2006 International Conference on* (pp. 447-453). IEEE. Doi: 10.1109/ICCES.2006.320489

Zhang, X. & Pham , H. (2000). An analysis of factors affecting software reliability, *Journal of Systems and Software*, Volume 50, Issue 1, Pages 43-56. Recuperado de <http://www.sciencedirect.com/science/article/pii/S0164121299000758>.

ANEXOS

Anexo A

La siguiente fue la encuesta que se realizó para obtener la información de los proyectos.

Encuesta de factores que inciden en el desarrollo de software

Con la presente encuesta se pretende recolectar información de proyectos de software que ya han sido terminados y cerrados, con el fin de analizar los datos y poder determinar factores de éxito en los proyectos.

Por cada proyecto de software en el que haya participado y se haya terminado por favor calificar los siguientes factores, las respuestas corresponden a su percepción de como considera que se desarrolló cada factor en el proyecto que va a evaluar.

Agradezco su atención y colaboración.

***Obligatorio**

Nombre del proyecto en el que participó *

Sección A

Califique de acuerdo a su percepción sobre cada factor del proyecto, donde: 1 Totalmente en desacuerdo, 2 Parcialmente en desacuerdo, 3 Indiferente, 4 Parcialmente de acuerdo, 5 Totalmente de acuerdo

Capacidad de los desarrolladores

Amplios conocimientos de los desarrolladores acerca de los procesos de desarrollo *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplio nivel de experiencia de los desarrolladores con la plataforma o entorno del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplio nivel de experiencia de los desarrolladores con las herramientas y el lenguaje de desarrollo del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplios conocimientos de los desarrolladores sobre las herramientas usadas para soportar actividades de desarrollo del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Conocimiento del dominio

Alta capacidad de los desarrolladores para entender la arquitectura y resolver los riesgos asociados del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplia experiencia previa de los desarrolladores con aplicaciones similares al proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplios conocimientos de los desarrolladores sobre los aspectos de negocio del producto a desarrollar *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplios conocimientos de los desarrolladores sobre el dominio para el que está destinado el producto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Amplios conocimientos de los desarrolladores sobre el producto a desarrollar *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Composición del equipo

Gran capacidad de comunicación de los desarrolladores *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Gran capacidad de los desarrolladores para resolver problemas *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Excelentes conocimientos técnicos de los desarrolladores *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Gran capacidad de comunicación de los diseñadores *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Gran capacidad de los diseñadores para resolver problemas *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Excelentes conocimientos técnicos de los diseñadores *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Complejidad del producto

Amplia complejidad de las interfaces externas del producto a desarrollar *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Madurez de Gestión de Proyectos

Excelente definición de roles y responsabilidades para las actividades de desarrollo del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Excelente definición de responsabilidades y niveles de autoridad para las actividades de desarrollo del proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Excelente cumplimiento con el plan de trabajo respecto a los procesos de la empresa *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Excelente cumplimiento con el plan de trabajo respecto al proyecto *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Ejecución del proyecto

Ejecución del proyecto dentro del tiempo esperado *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Ejecución del proyecto dentro del presupuesto esperado *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Ejecución del proyecto dentro del alcance esperado *

1 2 3 4 5

Totalmente en desacuerdo Totalmente de acuerdo

Sección B

Califique de acuerdo a la Indicación de cada pregunta

Composición del equipo

Nivel de experiencia en programación de los desarrolladores *

En años de experiencia, <=1 año en promedio: Calificación 1, >1<=2 años en promedio: Calificación 2, >2<=3 años en promedio: Calificación 3, >3<=4 años en promedio: Calificación 4, Superior a 5 años en promedio: Calificación 5

1 2 3 4 5

Nivel de experiencia del equipo en proyectos de desarrollo *

En años de experiencia, <= 1 año: Calificación 1, > 1<= 2 años: Calificación 2, > 2 <= 3 años: Calificación 3, > 3<= 4 años: Calificación 4, > 4 años: Calificación 5

1 2 3 4 5

Uso de personas con postgrado en el proyecto *

0 personas: Calificación 1, 1 persona: Calificación 2, 2 personas: Calificación 3, 3 personas: Calificación 4, Superior a 4 personas: Calificación 5

1 2 3 4 5

Tamaño del equipo del proyecto *

1-2 personas: Calificación 1, 3-4 personas: Calificación 2, 5-6 personas: Calificación 3, 7-8 personas:

Calificación 4, Superior a 9 personas: Calificación 5

1 2 3 4 5



Colaboración

Nivel de continuidad y estabilidad del equipo del proyecto *

1 proyecto: Calificación 1, 2 proyectos: Calificación 2, 3 proyectos: Calificación 3, 4 proyectos: Calificación 4, Superior a 5 proyectos: Calificación 5

1 2 3 4 5



Complejidad del producto

Cantidad de interfaces externas del producto (tanto de hardware como software) *

0 Interfaces: Calificación 1, 1 Interfaces: Calificación 2, 2 Interfaces: Calificación 3, 3 Interfaces: Calificación 4, Superior a 4 Interfaces: 5

1 2 3 4 5



Tamaño del producto a desarrollar *

Medido en casos de uso (CU), inferior a 2 CU: Calificación 1, Entre 3 y 4 CU: Calificación 2, entre 5 y 6 CU: Calificación 3, entre 7 y 8 CU: Calificación 4, Superior a 8 CU: Calificación 5

1 2 3 4 5



Enviar

Nunca envíes contraseñas a través de Formularios de Google.

100%: has terminado.

Con la tecnología de

Este contenido no ha sido creado ni aprobado por Google.
[Informar sobre abusos](#) - [Condiciones del servicio](#) - [Otros términos](#)