

efectivo del desempeño de los alumnos, identificando áreas que podrían requerir apoyo adicional. Esto lleva a intervenciones más oportunas y personalizadas para mejorar el rendimiento académico.

Actualización de contenidos:

La aplicación permite una fácil actualización de contenidos, crucial

en el campo de la programación, donde las tecnologías y las mejores prácticas evolucionan constantemente.

La aplicación de Aprendizaje Móvil en la asignatura "Lenguaje de Programación I" tiene el potencial de transformar la experiencia de aprendizaje. Proporciona flexibilidad, personalización y acceso

a recursos actualizados, contribuyendo a un mejor rendimiento académico y una comprensión más profunda de los conceptos de programación. Además, brinda a los docentes herramientas para un seguimiento más efectivo del progreso de los estudiantes, permitiendo intervenciones más oportunas y personalizadas.

Desarrollo de un sistema de información para predecir resultados probables del Sector Bancario dentro de la empresa Talentobank-RRHH Consultores, C.A., Estado Miranda. (período de diciembre 2017 a diciembre 2022 y proyección a un año)

Trabajo de Grado presentado como requisito parcial para optar al Título de Ingeniero en Informática
Jesús Gabriel Moreno Guevara

El presente Trabajo de Grado, tiene como objetivo general Desarrollar un sistema de información para predecir resultados probables del sector bancario dentro de la empresa Talentobank-RRHH Consultores, C.A., Estado Miranda. (período de diciembre 2017 a diciembre 2022 y proyección a un año). Se realizó una serie de interrogantes que dieron respuesta a los diferentes objetivos específicos y permitió realizar un diagnóstico de la situación actual para el diseño de dicho Sistema. El diseño aplicado a esta investigación se describe de tipo no experimental, descriptivo y de Campo, la población y muestra se representará por un total de (20) veinte colaboradores entre las diferentes direcciones de la empresa Talentobank-RRHH Consultores, C.A., la encuesta y el instrumento es el cuestionario para ambas variables, siendo este con preguntas dicotómicas, validado por (5) cinco expertos de diferentes áreas. Las metodologías utilizadas fueron: Ágil SCRUM XP y CRISP-DM, y el uso de herramientas como R, Python Flask, JavaScript, HTML5, CSS3, MongoDB, NoSQL, Frameworks Angular, Jupyter Notebook, Apache, AWS, GitHub y Workflows. Esta investigación tiene como finalidad generar un sistema adaptable a las necesidades de la empresa y con todos los controles de calidad y de seguridad para poder generar las proyecciones del Sector Bancario Venezolano.

Descriptores: Sistema, Machine Learning, Modelos Predictivos, Regresión, Sector Bancario, Control de Gestión, Metodología CAMEL, Procesos de Negocio, Aprendizaje Automático, Lenguaje en R, Lenguaje en Python, Base de Datos, Ingeniería en Informática, Minería de Datos, CRISP-DM, Talentobank-RRHH Consultores, C.A. y Universidad Alejandro de Humboldt.

Metodología:

La metodología usada en esta investigación fue la metodología XP, que se utiliza en el desarrollo de aplicaciones móviles cuando se trata de la flexibilidad, la adaptabilidad a cambios frecuentes y la entrega incremental de funcionalidades para satisfacer las necesidades cambiantes del usuario y del mercado.

Bases Teóricas:**Contenido Programático:**

Son los fundamentos conceptuales y metodológicos que respaldan la selección, organización y presentación de temas, conceptos y habilidades en un plan de estudios. Estas bases se apoyan en teorías pedagógicas y enfoques educativos específicos que buscan optimizar la enseñanza y el aprendizaje de los estudiantes en el contexto de la informática, asegurando la coherencia y relevancia del contenido.

Identificador de Fuentes:

Se refiere a los principios y conceptos que respaldan la información de nombres únicos a recursos de información. Esto se fundamenta en teorías de gestión de datos, que buscan garantizar la singularidad, accesibilidad y trazabilidad.

1 Introducción

El entorno empresarial actual está en constante evolución, lo que plantea desafíos para las empresas en términos de adaptación y competitividad. La gestión financiera adecuada se

vuelve crucial en este contexto, ya que permite tomar decisiones acertadas y alineadas con los objetivos corporativos. Sin embargo, la gestión financiera va más allá del manejo de recursos monetarios e implica la gestión de otros recursos no monetarios. Además, requiere una estrategia sólida y actualizada, así como una responsabilidad compartida en toda la organización. En este sentido, el análisis financiero se convierte en una herramienta fundamental para evaluar los resultados económicos y financieros de una empresa. Proporciona información sobre aspectos como liquidez, solvencia, endeudamiento, eficiencia, desempeño y rentabilidad, facilitando la toma de decisiones en la gestión empresarial. Sin embargo, para identificar problemas y realizar correcciones, es necesario contar con herramientas que permitan el análisis y la interpretación de los datos financieros.

En los últimos años, el uso de la inteligencia artificial (IA) y el aprendizaje automático ha experimentado un aumento significativo en el sector financiero. Estas tecnologías han demostrado ser valiosas para mejorar los procesos, reducir costos, detectar fraudes y gestionar el riesgo. El aprendizaje automático, en particular, ha sido importante en el desarrollo de modelos predictivos, ya que pueden aprender y mejorar de forma autónoma a medida

que se les proporciona nueva información. Con el objetivo de predecir los posibles resultados del Sector Bancario Venezolano, este estudio propone el desarrollo de un sistema de información con un modelo de riesgo basado en técnicas de aprendizaje automático. Se utilizarán los principales indicadores de riesgo según el método CAMEL, y se buscará determinar el nivel de riesgo adecuado para evaluar y crear variables apropiadas para el modelo.

Para lograrlo, se utilizarán los lenguajes de programación R y Python, que ofrecen una amplia gama de herramientas y librerías especializadas. Además, se aplicará una combinación de metodologías ágiles, como eXtreme Programming (XP), y la metodología estructurada Cross-Industry Standard Process for Data Mining (CRISP-DM). Esta combinación permite un enfoque ágil y organizado, garantizando la entrega temprana de valor y una respuesta eficiente a los desafíos del proyecto. En resumen, este estudio destaca la importancia de la gestión financiera adecuada en un entorno empresarial en constante evolución. El análisis financiero y el uso de técnicas de inteligencia artificial y aprendizaje automático son herramientas clave para la evaluación de riesgos en el Sector Bancario Venezolano. La combinación de los lenguajes de programación R y Python

con metodologías ágiles y estructuradas permite el desarrollo de un sistema de información efectivo y eficiente para la toma de decisiones financieras basadas en análisis de datos precisos y actualizados.

2 Planteamiento del Problema

Según Arias (2016) "el planteamiento del problema consiste en describir el problema de manera amplia la situación objeto de estudio, ubicándola en un contexto que permita comprender su origen, relaciones e incógnitas por responder" (p.41). El mundo avanza y con él, la digitalización; juntos van de la mano hacia una sociedad desarrollada y dependiente de la tecnología. En la era digital son necesarias nuevas técnicas que ayuden a las empresas a crecer y lograr un posicionamiento en su sector. De ahí lo relevante de prestar atención a los datos, ya que en la actualidad son muchos los que se genera y necesitan un orden, almacenaje y técnicas para sus análisis. En concreto, para lograr anticiparse a las necesidades y complejidades del mercado, es importante incorporar los análisis predictivos, a fin de aprovechar estos datos y desarrollar soluciones innovadoras que toda empresa necesita y consta de procesos que mediante técnicas computacionales, permiten anticiparse a las necesidades futuras y prevenir eventuales problemas.

Hoy en día, las empresas

son más conscientes sobre el impacto que tienen sus actividades de control sobre la información para la toma de decisiones, siendo una estrategia dentro del mercado, no solo para hacerse responsables de su consecución; sino también contribuir al crecimiento de la organización, y hacer frente a los desafíos en cuanto a la búsqueda de información para la formulación de indicadores clave de medición para conocer la situación en que se encuentra la misma.

Ahí reside la importancia de los modelos predictivos, que consiste en el uso de datos históricos destinados a la sondeo de patrones e identificación de tendencias que se pueden utilizar para reorganizar la forma mediante la cual se prestan servicios y desarrollar productos y el benchmarking se convierte en una herramienta apropiada para evaluar, alcanzar y conseguir optimizar los resultados para la empresa como lo indica Codling (1998) "benchmarking puede ser considerado como un proceso de evaluación de productos, servicios y procesos entre organizaciones, mediante el cual, una de ellas analiza cómo otra realiza una función específica para igualarla o mejorarla.", (p.34).

Además, los criterios para evaluar la eficiencia a través del benchmarking se pueden realizar con los indicadores clave financieros bajo la metodología camel,

la cual nace de la necesidad de medir el riesgo dentro de las instituciones del sector bancario y miden solidez, solvencia, gestión administrativa, rentabilidad que presentan las instituciones financieras. Según Bou-Said (2003) "la metodología CAMEL fue originalmente creada como una herramienta de análisis para ser usada por los entes encargados de la actividad reguladora en su función de analizar la condición financiera de los bancos e instituciones financieras..." (p.64).

En Talentobank-RRHH Consultores, C.A., se creó el proyecto piloto, creación y generación de valor para el desarrollo de soluciones tecnológicas y mejorar los resultados operacionales, con base al nuevo proyecto piloto, la empresa no cuenta con su propio sistema de modelos predictivos, por lo tanto, es una situación que le dificulta en la actualidad, al momento de presentar información que en ocasiones no coinciden los datos presentados lo que conlleva a confusiones o a conseguir los datos de manera atemporal y fragmentada. Otras compañías, cuentan en este momento con sus propios sistemas de modelos predictivos enmarcados dentro de sus estrategias de negocio, atendiendo a diferentes sectores y en particular el Sector Bancario Venezolano, sobre el comportamiento del sistema en general para la toma de decisiones que permite

mantener la rentabilidad y solidez, con la finalidad de valorar segmentos y posiblemente estrategias.

Con base a lo anteriormente expuesto, el presente estudio tiene como objetivo el desarrollo de un sistema de modelo predictivo que permita dar continuidad y contexto al control de la información del Sector Bancario Venezolano, de esta manera el desarrollo es de gran utilidad e importancia; ya que se pretende satisfacer las necesidades reales de información para la toma de decisión, clave y estratégica. Como los canales y procesos han tomado una gran importancia en ésta; y vital para contar con un sistema donde se pueda visualizar, controlar y monitorear permitiendo anticiparse a posibles problemas y corregir errores con un mayor margen de tiempo. Con este sistema la base de datos está poblada directamente de las cifras publicadas de la Superintendencia de las Instituciones del Sector Bancario (SUDEBAN).

2.1 Objetivo General

Desarrollar un sistema de información para predecir resultados probables del sector bancario dentro de la empresa Talentobank-RRHH Consultores, C.A., Estado Miranda. (período de diciembre 2017 a diciembre 2022 y proyección a un año).

2.2 Objetivos Específicos

Diagnosticar la situación actual de la empresa en cuanto a los requerimientos necesarios

para los análisis predictivos del Sector Bancario Venezolano.

Identificar las herramientas y técnicas de ingeniería de software que pueden ser aplicable al diseño de un sistema de información para los análisis predictivos del Sector Bancario Venezolano.

Establecer la metodología CAMEL y los indicadores financieros claves que pueden ser aplicables al diseño de un sistema de información para los análisis predictivos del Sector Bancario Venezolano.

Proponer un sistema factible, dentro de la empresa, flexible para mejorar los análisis predictivos del Sector Bancario Venezolano.

3 Identificación, Definición y Operacionalización de las Variables

3.1 Tabla #1 Identificación y Definición de las Variables.

Objetivos Específicos / Variables / Definición Conceptual de las Variables.

1. Diagnosticar la situación actual de la empresa en cuanto a los requerimientos necesarios para los análisis predictivos del Sector Bancario Venezolano. Requerimientos necesarios para realizar los Análisis Predictivos del Sector Bancario Venezolano.

2. Identificar las herramientas y técnicas de ingeniería de software que pueden ser aplicable al diseño de un sistema de información para los análisis predictivos del Sector Bancario Venezolano.

Herramientas y Técnicas de Ingeniería de Software.

3. Establecer la metodología CAMEL y los indicadores financieros claves que pueden ser aplicables al diseño de un sistema de información para los análisis predictivos del Sector Bancario Venezolano. Técnicas de Análisis de Eficiencia Financiera.

4. Proponer un sistema factible, dentro de la empresa, flexible para mejorar los análisis predictivos del Sector Bancario Venezolano. Sistema de Gestión.

3.2 Tabla #2 Operacionalización de las Variables.

Variables / Dimensión / Indicador / Instrumento / Ítems.

1. Requerimientos necesarios para realizar los análisis predictivos del Sector Bancario Venezolano. Gerencial. 1. Disponibilidad de la información. 2. Disponibilidad de las herramientas tecnológicas. Cuestionario. 1, 2,.

2. Herramientas y Técnicas de Ingeniería de Software. Arquitectura de datos y Seguridad de datos. 3. Modelos de datos conceptuales, lógicos y físicos. 4. Ciclo de vida de los datos. 5. Nivel de permisología y acceso a usuarios. 6. Nivel de permisología, protección y seguridad de datos. Cuestionario. 3, 4, 5, 6,.

3. Técnicas de Análisis de Eficiencia Financiera. Indicadores y Modelo de Riesgo. 7. Diseño y configuración general de

los Modelos de Machine Learning con la Metodología CAMEL. Cuestionario. 7.

4. Sistema de Gestión. Tecnología. 8. Lenguaje de programación. 9. Bases de Datos. 10. Interfaz gráfica (PHP). Cuestionario. 8, 9, 10.

4 Metodología y Aplicación

4.1 Diseño y Tipo de Investigación

El diseño aplicado a esta investigación se describe de tipo no experimental, descriptivo y de Campo, la encuesta y el instrumento es el cuestionario para ambas variables, siendo este con preguntas dicotómicas, validado por (5) cinco expertos de diferentes áreas.

4.2 Población y Muestra

La población y muestra se representará por un total de (20) veinte colaboradores entre las diferentes direcciones de la empresa Talentobank-RRHH Consultores, C.A.

Talentobank-RRHH Consultores, C.A.	
Estructura Organización	Posiciones
Dirección de Sistemas y TI	Directores
	Especialistas
	Sub-Total 6
Dirección de Desarrollo de Negocios	Directores
	Gerentes
	Analistas
Sub-Total 9	
Dirección Académica	Directores
	Consultores
Sub-Total 2	
Dirección de Operaciones y Finanzas	Directores
	Especialistas
Sub-Total 3	
# Colaboradores 20	

Tabla 1 - Muestra Estructura Organizacional 2023

Talentobank-RRHH Consultores, C.A.	
Estructura Organización	# Colaboradores
Dirección de Sistemas y TI	6
Dirección de Desarrollo de Negocios	9
Dirección Académica	2
Dirección de Operaciones y Finanzas	3
20	

Tabla 2 - Población Estructura Organizacional

4.3 Técnicas e Instrumentos de Recolección de Información

Consiste en la aplicación de métodos estadísticos a los resultados obtenidos de la aplicación del instrumento de evaluación, para posteriormente interpretarlos, así como lo indica el Manual de Trabajo de Grado de la Universidad Alejandro de Humboldt (2016), "Se refiere a las distintas técnicas que pueden utilizarse con el

frecuencia, moda, mediana, varianza, desviación típica, entre otros, lo que facilitó la agrupación y análisis de los datos obtenidos.

Luego de definir la técnica y el instrumento de recolección de datos, se optó por emplear como técnica la encuesta y como instrumento el cuestionario. El cuestionario para utilizar fue de tipo dicotómico en donde se realizó 10 preguntas con respuesta de "Sí" y "No".

Las preguntas a

Panel de Especialistas (Docentes y Empresarial)		
Profesor Univ.	Guzmán, Elisaira	Ingeniería
Profesor Univ.	De Faria, Alice	Metodología
Profesor Univ.	Fortunato, Tony	Metodología
Profesor Univ.	Yáñez, Horacio	Estadística
Tutor Emp.	Díaz, Athena	Empresarial

Tabla 3 - Juicio de Expertos Académicos y Empresarial

fin de interpretar los datos obtenidos; de esta manera, se puede contrastar las opiniones de los sectores involucrados en el estudio con la teoría utilizada". (p.33). En base a lo anteriormente descrito, luego de recopilar toda la información necesaria, se procedió a aplicar a la población de veinte (20), el instrumento de recolección de datos, que para efectos de este trabajo de investigación fue la encuesta, a la cual, posteriormente de ser aplicada y registrada de forma manual, se hizo uso de la estadística descriptiva como técnica de análisis de datos, la cual está compuesta por:

continuación:

1. ¿Usted considera que la información que está disponible en la empresa es suficiente para el sistema de gestión?

2. ¿Acerca de las herramientas tecnológicas con las que cuenta la empresa cumple sus expectativas?

3. ¿Usted evidencia que no cuenta con los recursos de modelos de datos conceptuales necesarios?

4. ¿Estaría usted de acuerdo en la mejora de la secuencia de etapas por las que pasan los datos en su ciclo de vida?

5. ¿Usted considera que los niveles de permisología y acceso a usuarios cumplen

con todos los protocolos de seguridad de la información?

6. ¿Usted evidencia que los niveles de permisología, protección y seguridad de datos son los adecuados?

7. ¿El diseño y configuración general de los modelos de machine learning con la metodología CAMEL que provee es clara?

8. ¿Considera usted que los lenguajes de programación utilizados son los adecuados para generar la información que necesita?

9. ¿Evidencia usted que la estructura y los diseños de bases de datos son los ajustados para resguardar la información utilizada?

10. ¿En general, estuvo satisfecho con el uso del sistema de control de gestión?

4.4 Diseño Metodológico

Las perspectivas que aquí se adoptan e involucran el desarrollo de un sistema de software y la exploración de datos es la combinación de dos metodologías que abarca la Minería de Datos y un entorno Ágil, por lo tanto, Cross-Industry Standard Process for Data Mining (CRISP-DM) para la minería de datos, que describe ciertas actividades del tipo de proyecto. Sin embargo, en este caso, se muestra específicamente las actividades para un proyecto factible de analítica de datos en un entorno ágil para el desarrollo del sistema. Las actividades se dividen en dos tipos: de producto y de proceso. Las actividades de producto describen lo que se debe hacer para construir los entregables en cada fase,

mientras que las actividades de proceso se realizan en paralelo para controlar los recursos asignados. Luego, se presenta una selección de actividades ajustadas a las necesidades particulares del proyecto factible, incluyendo actividades de producto y algunas de proceso de eXtreme Programming. Finalmente, se muestra una lista de actividades finales para CRISP-DM Ágil después de cada punto.

Fase I: Comprensión del Negocio: en esta fase, se entiende la problemática con el levantamiento de información de la situación actual de la empresa, identificando los objetivos de la investigación, determinando los recursos disponibles, y estableciendo un plan preliminar. La visión ágil incluye lo siguiente: determinar objetivos del negocio, evaluar la situación, determinar objetivos de minería de datos, realizar plan de proyecto y plan de proyecto.

Metodología CRISP-DM y XP fase I

Determinar objetivos del negocio: se definen los objetivos generales y/o específicos para el proyecto de analítica.

Evaluar la situación: implica una búsqueda de datos más detallada sobre todos los recursos, restricciones, suposiciones y otros factores que deben considerarse al determinar el objetivo del análisis de datos y el plan del proyecto.

Individuos	Preguntas										
	p1	p2	p3	p4	p5	p6	p7	p8	p9	p10	
1	1	1	1	1	1	1	1	1	1	1	10
2	1	1	1	1	1	1	1	1	1	1	10
3	1	1	1	1	1	1	1	1	1	1	10
4	1	1	1	1	1	1	1	1	1	1	10
5	1	1	1	1	1	1	1	1	1	1	10
6	1	1	1	0	1	1	1	1	1	1	9
7	1	1	0	0	1	1	1	1	1	1	8
8	1	1	1	0	1	1	1	1	1	1	9
9	1	1	1	0	1	1	1	1	1	1	9
10	1	1	1	1	1	1	1	1	1	1	10
11	1	1	1	1	0	1	1	1	1	1	9
12	1	0	0	0	1	1	1	1	1	1	7
13	1	1	1	1	1	1	1	1	1	1	10
14	0	1	1	0	1	1	1	1	1	1	8
15	1	1	1	0	1	1	1	1	1	1	9
16	1	1	0	0	1	0	1	1	1	1	7
17	1	1	0	1	0	0	1	1	1	1	7
18	0	0	0	1	0	0	1	0	0	1	3
19	0	0	0	1	0	0	0	0	0	0	1
20	0	1	0	0	0	0	0	0	0	0	1
Totales	16	17	13	11	15	15	18	17	17	18	
p	0,80	0,85	0,65	0,55	0,75	0,75	0,90	0,85	0,85	0,90	
q	0,20	0,15	0,35	0,45	0,25	0,25	0,10	0,15	0,15	0,10	
p*q	0,16	0,13	0,23	0,25	0,19	0,19	0,09	0,13	0,13	0,09	
Σ(p*q)	1,5725										
e2	8,3447										
k	10										

Tabla 4 - Resultados de la encuesta

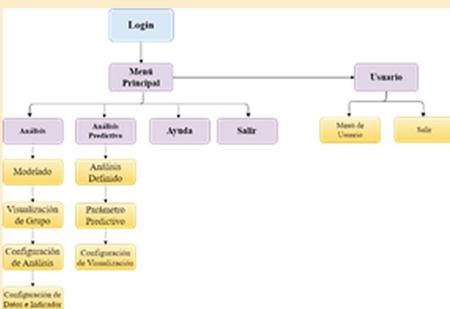
Determinar objetivos de minería de datos: un objetivo de negocio establece objetivos en la terminología de negocios. Un objetivo de minería de datos establece los objetivos del proyecto en términos técnicos.

Realizar plan de proyecto: incluye las siguientes actividades Crear Historias de Usuario.

Plan de proyecto: definir un backlog de producto con todas las actividades a realizar de forma organizada y priorizada.



Ilustraciones 1 -Actividades de la Metodología CRISP-DM



Ilustraciones 2 -Modelo del Sistema Propuesto Ampliado y XP fase I

Fase II: Comprensión de los Datos: en esta fase, la empresa se encargó de reunir los datos necesarios para el proyecto y asegurarse de que estén estructurados. Esto puede implicar el uso

de herramientas de minería y limpieza de datos para asegurarse de que los datos sean precisos y útiles para el análisis. Una vez que los datos estén recopilados, el siguiente paso es explorarlos, es decir, Esto implica el uso de técnicas de análisis exploratorio de datos (EDA) para identificar patrones, distribuciones, valores atípicos, relaciones y correlaciones entre las diferentes variables, el analizar la correlación entre las diferentes variables de los datos, ayuda a identificar qué variables son importantes y cuáles no lo son. La visión ágil incluye lo siguiente: crear backlog de sprint, recolectar y describir los datos, verificar la calidad de los datos, seleccionar los datos y listar indicadores de negocio.

Metodología CRISP-DM fase II

Crear backlog de sprint: según la capacidad del equipo de trabajo, se listan las historias de usuario priorizadas que fue ejecutada en la próxima iteración o sprint.

Recolectar y describir los datos: se adquieren los datos (o acceso a los datos) enumerados en los recursos del proyecto, se examinó las propiedades "brutas" o "superficiales" de los datos adquiridos y se informa sobre los resultados.

Verificar la calidad de los datos: se examinó la calidad de los datos, abordando preguntas como: ¿Están completos los datos (cubren todos los casos

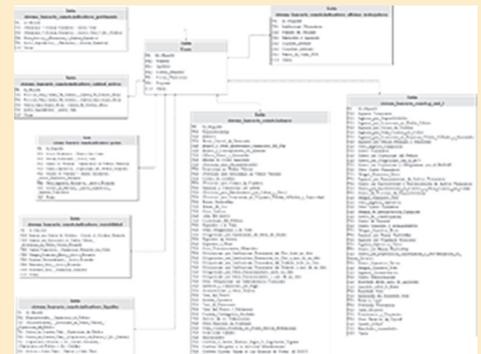
requeridos)? ¿Es correcto o contiene errores y, si hay errores, qué tan comunes son? ¿Hay valores faltantes en los datos? Si es así, ¿cómo se representan, ¿dónde se producen y qué tan comunes son?

Seleccionar los datos: decidir sobre los datos que se utilizó para el análisis. Los criterios incluyen relevancia para los objetivos de minería de datos, calidad y restricciones técnicas, como los límites en el volumen de datos o los tipos de datos.

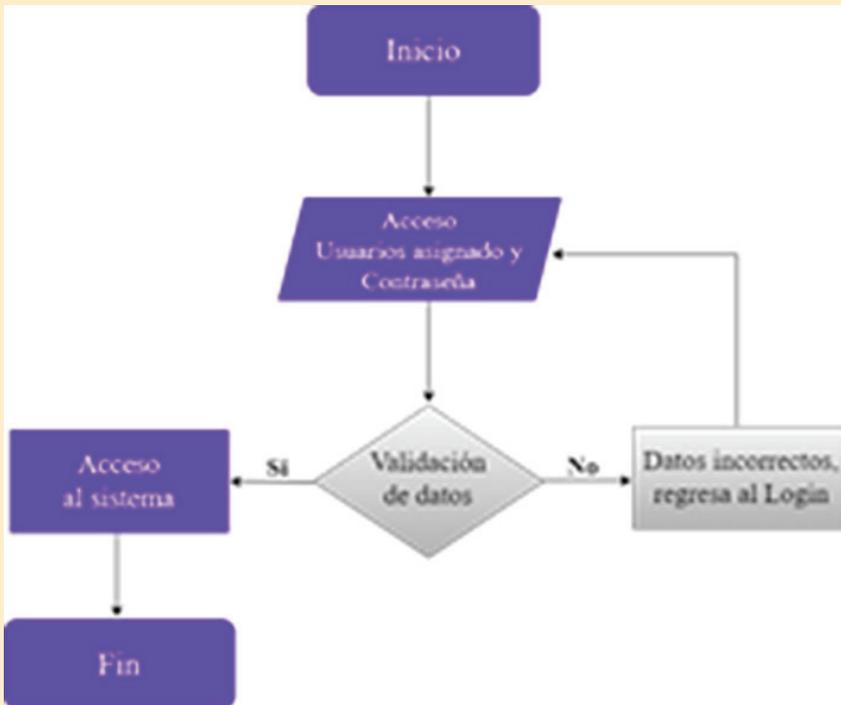
Listar indicadores de negocio: basado en los objetivos de minería de datos, se relacionan los indicadores CAMEL a utilizar para la visualización en los diferentes visores del sistema.



Ilustraciones 3 -Actividades de la Metodología CRISP-DM y XP fase II



Ilustraciones 4 -Diagrama Entidad-Relación



Ilustraciones 5 -Diagrama de Flujo de Datos de Acceso (Inicio)

Para comprender el diagrama de entidad-relación se realizó el siguiente análisis: en todas las tablas que corresponden al sistema, están relacionadas entre sí, con un valor mínimo de uno a uno, las cuales tienen la función de validaciones al momento de realizar la programación planteada y guardar los registros sin tener una redundancia de datos dentro de las mismas tablas del sistema.

Fase III: Preparación de los Datos: en esta fase, en primer lugar, la recopilación y calidad de los datos fue necesario para el proyecto de modelos predictivos, con la finalidad del preprocesamiento de los datos que incluirán una serie de tareas, como la eliminación de ruido, la normalización de los datos, la selección de características relevantes y la transformación de los datos. Estas tareas fueron

necesarias para garantizar que los datos sean adecuados para el análisis y la modelización. Además, dividir los datos en conjuntos de entrenamiento y prueba. Una vez que los datos están preparados, fue necesario seleccionar el modelo que se utilizó y ajustó los parámetros del modelo. Esto implica la evaluación de varios modelos y la selección del modelo que mejor se ajuste a los datos. Finalmente, es necesario evaluar la precisión del modelo en los datos de prueba. La visión ágil incluye lo siguiente: limpiar los datos, construir e integrar los datos, construir mock-ups y realizar sprint actual.

Metodología CRISP-DM fase III

Limpiar los datos: se aumentó la calidad de los datos al nivel requerido por las técnicas de análisis seleccionadas. Esto puede implicar la selección de

subconjuntos limpios de los datos, la inserción de valores predeterminados adecuados o técnicas más ambiciosas, como la estimación de los datos faltantes mediante el modelado.

Construir e integrar los datos: estos son métodos mediante los cuales la información se combina de varias tablas o registros para crear nuevos registros o valores.

Construir mock-ups: se realiza una vista previa del visor para ser evaluado por el usuario final antes de su implementación.

Realizar sprint actual: el equipo de trabajo se encarga de realizar las tareas asignadas en el sprint backlog.

Fase IV: Modelado: en esta fase, la empresa se encargó de la selección del algoritmo adecuado depende del problema y de los datos que se tienen. Se deben evaluar varios algoritmos para determinar cuál es el más adecuado para el problema. Una vez seleccionado el algoritmo, se construye el modelo utilizando los datos de entrenamiento. Es importante



Ilustraciones 6 -Actividades de la Metodología CRISP-DM y XP fase III

ajustar los parámetros del modelo para obtener los mejores resultados. Se evalúa el modelo utilizando los datos de prueba para medir su precisión y su capacidad para generalizar a nuevos datos. Finalmente, se implementa el modelo en un sistema o aplicación para que pueda ser utilizado en la empresa. Es importante seguir monitorizando el modelo y actualizándolo regularmente para asegurarse de que siga siendo preciso y relevante. La visión ágil incluye lo siguiente: crear esquema de visualización, seleccionar técnicas de presentación de datos, construir los tableros de control, evaluar los tableros de control y realizar Daily Stand-Up.

Metodología CRISP-DM fase

IV

Crear esquema de visualización: basado en los datos y las técnicas seleccionadas, se procede a especificar los indicadores a usar en los visores del sistema, según los entregables definidos por el sprint.

Análisis de indicadores: lista de los indicadores (CAMEL), con la explicación de cada uno. Es importante destacar la naturaleza del indicador, el origen de los datos necesarios para calcularlo, la forma en la que se realizó el cálculo y la meta que se pretende alcanzar inicialmente para el indicador.

Análisis de reportes: lista de los reportes a realizar con la explicación de cada uno. Es importante destacar el tipo de reporte (interfaz, matriz, rompimiento, etc.), el origen de los datos necesarios para

generarlo y cómo se puede realizar el drill down, técnica usada para ampliar el detalle de los elementos presentados.

Diseño de reportes (reporting): imágenes de los prototipos de los reportes ilustrando en lo posible datos reales o aproximados de la situación actual y la forma de ampliar el detalle de algunos elementos del reporte.

Diseño de tablero: (visores): imagen del prototipo del tablero de control ilustrando en lo posible datos reales o aproximados de la situación actual y la forma de ampliar el detalle de algunos elementos del tablero de control.

Seleccionar técnicas de presentación de datos: como primer paso en el modelado, se seleccionó la técnica de presentación real que se va a utilizar. Aunque es posible que ya haya seleccionada una herramienta durante la fase de 'entendimiento de negocio', esta tarea se refiere a la técnica de presentación específica. Si se aplican múltiples técnicas, se realiza esta tarea por separado para cada técnica.

Construir los tableros de control: se construyó el esquema de visualización basado en el mock-up aprobado por el usuario, para mostrar los datos creados previamente.

Evaluar los tableros de control: se socializó el producto con el cliente para que tenga interacción con el entregable y pueda retroalimentar según lo vea necesario de acuerdo con su juicio de expertos como dueños del negocio.

Realizar Daily Stand-

Up: reunión diaria para el seguimiento del equipo de trabajo y actualización del estado de las actividades en el Kanban.

Fase V: Evaluación: en esta fase, se evaluó la precisión del modelo utilizando los datos de prueba para medir su capacidad para predecir los resultados correctos. Es importante tener en cuenta que la precisión no fue la única métrica importante, ya que un modelo con una precisión alta puede no ser adecuado para el uso previsto. Fue importante evaluar la fiabilidad del modelo, lo que significa determinar si el modelo fue coherente en diferentes conjuntos de datos y si su precisión se mantiene en el tiempo. Fue necesario evaluar la relevancia del modelo y asegurarse sea el adecuado para el problema específico que se estará tratando de resolver y que tenga una precisión suficiente para ser útil y ayudó a identificar posibles limitaciones del modelo y a determinar si



Ilustraciones 7 -Actividades de la Metodología CRISP-DM y XP fase IV

necesita mejoras o ajustes adicionales. La visión ágil incluye lo siguiente: evaluar los resultados, revisión de sprint y determinar próximos pasos.

Metodología CRISP-DM fase

V

Evaluar los resultados: se socializó el entregable con todos los stakeholders para determinar la veracidad de la información suministrada y adaptarse a los objetivos definidos en la fase 'Entendimiento de los Datos'.

Revisión de sprint: se realizó a nivel de proceso una entrega formal del entregable mediante una ceremonia de socialización.

Determinar próximos pasos: dependiendo de los resultados de la evaluación y la revisión del proceso, el equipo del proyecto decide cómo proceder. El equipo decide si terminar el proyecto y continuar con la implementación, iniciar nuevas iteraciones o configurar nuevos proyectos de minería de datos.

Fase VI: Despliegue: en esta fase, fue necesario integrar el modelo en el sistema de producción de la empresa. Esto puede implicar la integración del modelo en una aplicación o en un sistema existente, realizar pruebas rigurosas del modelo antes de su implementación. Esto incluye pruebas de integración, de unidad y de

aceptación para asegurarse que el modelo funcione correctamente. También, fue necesario configurar el entorno de producción, servidores, bases de datos y otros componentes necesarios para el modelo funcione adecuadamente. Reentrenamiento del modelo con nuevos datos o la mejora de los algoritmos utilizados, la comunicación de los resultados del modelo a los interesados en la empresa. Esto les ayudó a entender cómo el modelo estuvo afectando a la empresa y cómo se está utilizando éstos; para tomar decisiones informadas. El llevar a cabo un mantenimiento regular del modelo para asegurarse que sigue funcionando correctamente es importante



Ilustraciones 7 -Actividades de la Metodología CRISP-DM y XP fase IV

y que cumple con los requisitos de precisión y fiabilidad. La visión ágil incluye lo siguiente: despliegue del plan y retrospectiva de sprint

A modo de resumen, la Figura 1 muestra las fases de un proyecto factible de combinación de

metodologías.

Metodología CRISP-DM fase

VI

Despliegue del plan: esta tarea toma los resultados de la evaluación y determina una estrategia para el despliegue. Si se identificó un procedimiento general para crear los modelos relevantes, este procedimiento se documenta en el backlog de producto para su posterior implementación.

Retrospectiva de sprint: fue una oportunidad para que el equipo se inspeccione a sí mismo y cree un plan para que se realicen mejoras durante el próximo Sprint. El propósito de la retrospectiva de sprint es:

Inspeccionar cómo fue el último sprint con respecto a las personas, las relaciones, el proceso y las herramientas.

Identificar y ordenar los elementos principales que salieron bien y las mejoras potenciales.

Crear un plan para implementar mejoras en la forma en que el equipo hace su trabajo.

5 Resultados

El proyecto de desarrollar un sistema de información para predecir resultados probables del sector bancario dentro de la empresa Talentobank-RRHH Consultores, C.A., tiene varios

beneficiarios e impactos, entre ellos:

Este proyecto beneficia directamente a la empresa al proporcionarle herramientas y conocimientos para tomar decisiones informadas basadas en la predicción de resultados probables del sector bancario. De esta manera, la empresa puede mejorar su eficiencia y eficacia en la gestión de recursos, lo que puede llevar a un aumento en su rentabilidad y competitividad.

Los clientes de la empresa también se benefician indirectamente de este proyecto, ya que la empresa puede proporcionarles un mejor servicio y asesoramiento basado en la información que se ha recopilado y analizado a través del sistema de información.

Los empleados de la empresa también se benefician del proyecto, ya que les proporciona la oportunidad de mejorar sus habilidades y conocimientos en el análisis de datos y la predicción de resultados del sector bancario.



Ilustraciones 9 -Actividades de la Metodología CRISP-DM y XP fase VI

proyecto fue significativo en términos de la mejora de la eficiencia y eficacia de la empresa en la gestión de sus recursos y en la mejora del servicio y asesoramiento proporcionado a sus clientes. Además, el proyecto puede contribuir a la innovación en el sector bancario y a la mejora del desempeño económico de la empresa.

6 Conclusiones

Luego de haber culminado el proceso de investigación orientado hacia el desarrollo de un sistema de información para predecir resultados probables del sector bancario dentro de la empresa Talentobank-RRHH Consultores, C.A., Estado Miranda. (período de diciembre 2017 a diciembre 2022 y proyección a un año).

Se diagnosticó las mejoras necesarias en la recopilación, almacenamiento y procesamiento de datos, así como la capacitación del personal en análisis predictivo. Se destacó la importancia de utilizar herramientas y técnicas de ingeniería de software, como el ciclo de vida del software, metodologías ágiles y el uso de R, Python, minería de datos y algoritmos de aprendizaje automático. Se estableció la metodología CAMEL y los indicadores financieros clave para una gestión de riesgos más eficiente. Se propone un sistema factible, dentro

de la empresa para mejorar los análisis predictivos del Sector Bancario Venezolano, en donde fue necesario realizar pruebas y ajustes para su mejora continua, así como involucrar a los usuarios finales en el proceso de diseño y pruebas para asegurar su eficacia y eficiencia, y se enfatizó en aumentar la eficiencia operativa, generar ahorros de costos y mejorar la satisfacción del cliente mediante la innovación tecnológica.

Las recomendaciones a partir de las conclusiones previas, se recomienda a Talentobank-RRHH Consultores, C.A. tomar en cuenta ciertos aspectos antes de implementar soluciones de análisis predictivo:

Talentobank-RRHH Consultores, C.A., en primer lugar, es importante que la empresa para mantener la operatividad del sistema y su precisión debe realizar una revisión detallada de los datos disponibles, tanto internos como externos, para determinar su calidad y relevancia para el análisis predictivo.

Se recomienda, complementar con la incorporación de otras técnicas de Minería de Datos más avanzadas como Machine Learning o Deep Learning que pueden ser ejecutadas para potenciar

los análisis de riesgos para la toma de decisiones.

Se sugiere, realizar mantenimiento y actualización de hardware y software preventivo a los equipos que almacenan y conforman la base de datos y el sistema de la empresa.

Asimismo, se recomienda realizar periódicamente capacitaciones en general a los colaboradores encargado de los procesos y funciones del sistema.

Por último, programar backup de la data disponible almacenada en el servidor centralizado, de modo que se evite pérdida de la información.

7 Referencias

Arias, F. (2006). El proyecto de Investigación. Introducción a la Metodología Científica. 5a Edición. Caracas: Episteme, C.A.

Abad, S. (2006). Adquisición e Implementación de Sistemas Informáticos. Caracas: Contexto Ediciones.

Alfonzo, I. (1994). Técnicas de Investigación Bibliográfica. Caracas: Contexto Ediciones.

Baca, G. (2006). Formulación y Evaluación de Proyectos Informáticos. México: McGraw-Hill Interamericana.

Balestrini, M. (2006).

Como se Elabora el Proyecto de Investigación (7ma ed.). Caracas: BL Consultores Asociados, C.A.

De Oliveira, R. A., Ferreira, R. B., & Da Silva, J. E. (2010). Combining XP and CRISP-DM for data warehouse development: a case study. In Proceedings of the 2010 ACM Symposium on Applied Computing (p. 1958-1963).

García-Sánchez, F., González-Rodríguez, M., & Colomo-Palacios, R. (2016). Integrating CRISP-DM and XP for data mining projects in small and medium-sized enterprises. *Journal of Systems and Software*, 114, p. 83-96.

Gonzalo Bello R. (2009). Operaciones Bancarias en Venezuela, Teoría y práctica. Universidad Católica Andrés Bello. Editorial Texto, C.A.

Guajardo G. y Andrade De Guajardo N. (2008). Contabilidad Financiera. Quinta edición. McGraw-Hill Interamericana.

Ortiz A, Héctor, Análisis Financiero Aplicado y Principios de Administración Financiera. 2008, 13.a Edición.

Sánchez, J. (2018). Sistema de Gestión de Información de la Universidad de Carabobo para el Ranking, Utilizando Arquitectura de Microservicios. Caso de estudio: información de estudiantes de pregrado y posgrado. p. 46.

Wild J., Subramanyam K. y Halsey R. (2007). Análisis de Estados Financieros. Novena edición. McGraw-Hill Interamericana.

Yang, Y., Goh, T. N., & Loh, H. T. (2007). A hybrid methodology of CRISP-DM and extreme programming for web application development. *Expert Systems with Applications*, 32(4), p. 1124-1134.

Fuentes Electrónicas
¿Qué es un sistema de gestión y para qué sirve? [Página Web en Línea]. Disponible en: <https://isbl.eu/2021/01/que-es-un-sistema-de-gestion-y-para-que-sirve/>. [Consulta Enero 2023].

¿Qué es un diagrama de flujo? [Página Web en Línea]. Disponible en: <https://www.lucidchart.com/pages/es/que-es-un-diagrama-de-flujo>. [Consulta Diciembre 2022].

La Importancia del Machine Learning [en línea]. Disponible en: <https://p300.net/p300-blog/la-importancia-del-machine-learning-becurious/> [Consulta Enero 2023].

El Marco Arquitectónico Big Data para el Tratamiento de Datos [en línea]. Disponible en: <https://blog.powerdata.es/el-valor-de-la-gestion-dedatos/tratamiento-dedatos-de-big-data-en-en-el-internet-de-las-cosas> [Consulta Enero 2023].