

PROPUESTA DE TRABAJO DE GRADO

PROGRAMA DE INGENIERIA DE SISTEMAS

DISEÑO E IMPLEMENTACIÓN DE UN SISTEMA WEB DE ANÁLISIS FINANCIERO CON MODELO DE MACHINE LEARNING PARA LA EVALUACIÓN DE RIESGO EN PYMES DEL MUNICIPIO DE IBAGUÉ, TOLIMA

Presentado por:

[NOMBRE COMPLETO DEL ESTUDIANTE]

[CODIGO ESTUDIANTIL]

Director propuesto:

[NOMBRE DEL DIRECTOR]

Línea de investigación:

Sistemas de Información, Desarrollo de Software e Inteligencia Artificial

Modalidad:

Desarrollo de Software / Proyecto de Investigación Aplicada

[NOMBRE DE LA UNIVERSIDAD]

FACULTAD DE INGENIERIA

PROGRAMA DE INGENIERIA DE SISTEMAS

Ibague, Tolima - FEBRUARY 2026

CONTENIDO

| | |
|---|----|
| 1. INFORMACION GENERAL DEL PROYECTO | 3 |
| 2. PLANTEAMIENTO DEL PROBLEMA | 4 |
| 3. JUSTIFICACION | 5 |
| 4. OBJETIVOS | 6 |
| 5. ALCANCE Y DELIMITACION | 7 |
| 6. MARCO TEORICO PRELIMINAR | 8 |
| 7. FUENTES DE DATOS | 9 |
| 8. ARQUITECTURA PROPUESTA | 10 |
| 9. METODOLOGIA | 11 |
| 10. CRONOGRAMA | 12 |
| 11. PRESUPUESTO | 13 |
| 12. RESULTADOS ESPERADOS | 13 |
| 13. REFERENCIAS BIBLIOGRAFICAS | 14 |

1. INFORMACION GENERAL DEL PROYECTO

| Campo | Descripcion |
|-------------------------|---|
| Titulo | Diseno e Implementacion de un Sistema Web de Analisis Financiero con Modelo de Machine Learning para la Evaluacion de Riesgo en PYMES del Municipio de Ibague, Tolima |
| Area de conocimiento | Ingenieria de Software, Inteligencia Artificial, Sistemas de Informacion |
| Linea de investigacion | Desarrollo de Software y Ciencia de Datos |
| Tipo de proyecto | Desarrollo tecnologico / Investigacion aplicada |
| Duracion estimada | 16 semanas (4 meses) |
| Delimitacion geografica | Municipio de Ibague, Departamento del Tolima, Colombia |
| Poblacion objetivo | Pequenas y Medianas Empresas (PYMES) de Ibague |

Resumen Ejecutivo

El presente proyecto propone el diseno e implementacion de un sistema web que permita a las PYMES del municipio de Ibague realizar analisis automatizado de sus estados financieros, calcular indicadores clave de desempeno y obtener una evaluacion de riesgo financiero basada en un modelo de Machine Learning. El sistema utilizara datos reales provenientes de dos fuentes oficiales: el Sistema de Informacion y Reporte Empresarial (SIREM) de la Superintendencia de Sociedades y la base de datos de empresas activas de la Camara de Comercio de Ibague, realizando un cruce de informacion mediante el NIT para identificar especificamente las empresas ibaguenenas.

2. PLANTEAMIENTO DEL PROBLEMA

2.1 Descripcion de la situacion problematica

Las Pequeñas y Medianas Empresas (PYMES) constituyen el motor económico de Ibagué y del departamento del Tolima. Según datos de la Cámara de Comercio de Ibagué, la ciudad cuenta con miles de empresas registradas, de las cuales la gran mayoría corresponde a micro, pequeñas y medianas empresas dedicadas principalmente a actividades comerciales, de servicios y agroindustriales.

Sin embargo, estas empresas enfrentan una problemática crítica: la deficiente gestión financiera derivada de la falta de herramientas tecnológicas accesibles. Las soluciones empresariales disponibles en el mercado (SAP, Oracle Financials, Microsoft Dynamics) tienen costos de licenciamiento que oscilan entre \$30,000 y \$100,000 USD anuales, cifras completamente inviables para pequeñas empresas con recursos limitados.

Esta situación genera que los empresarios ibaguenenses tomen decisiones financieras basadas en intuición o información incompleta, sin comprender indicadores fundamentales como ratios de liquidez, rentabilidad, endeudamiento o eficiencia operativa. El resultado es una alta tasa de mortalidad empresarial: estadísticas nacionales indican que aproximadamente el 70% de las PYMES colombianas fracasan en sus primeros cinco años de operación, siendo la deficiente gestión financiera una de las causas principales.

PROBLEMA CENTRAL IDENTIFICADO:

Las PYMES del municipio de Ibagué carecen de herramientas tecnológicas accesibles que les permitan: (1) analizar sistemáticamente sus estados financieros bajo normas NIIF, (2) calcular e interpretar indicadores financieros clave, y (3) evaluar su nivel de riesgo financiero para tomar decisiones informadas que mejoren su sostenibilidad.

2.2 Formulacion del problema

¿Como diseñar e implementar un sistema de información web que, mediante el análisis automatizado de estados financieros y la aplicación de técnicas de Machine Learning, permita a las PYMES del municipio de Ibagué evaluar su situación financiera y nivel de riesgo para apoyar la toma de decisiones gerenciales?

2.3 Sistematizacion del problema

- ¿Cuáles son los indicadores financieros más relevantes para evaluar la salud financiera de una PYME bajo normas NIIF?
- ¿Cómo integrar datos de múltiples fuentes (SIREM y Cámara de Comercio) para obtener información financiera específica de empresas de Ibagué?
- ¿Qué modelo de Machine Learning es más adecuado para clasificar el nivel de riesgo financiero de una PYME?
- ¿Cómo diseñar una interfaz web intuitiva que permita a usuarios no técnicos interpretar indicadores financieros complejos?
- ¿Qué métricas de usabilidad y precisión debe cumplir el sistema para ser útil a los empresarios ibaguenenses?

3. JUSTIFICACION

3.1 Justificacion economica y social

El Tolima es un departamento con vocacion agricola, comercial y de servicios, donde las PYMES representan la base del desarrollo economico regional. Ibague, como capital, concentra la mayor actividad empresarial del departamento. Proporcionar una herramienta de analisis financiero con capacidad predictiva a estas empresas contribuiria directamente a reducir la tasa de mortalidad empresarial, preservar empleos y fortalecer el tejido economico local.

3.2 Justificacion por disponibilidad de datos

El proyecto es viable gracias a la existencia de dos fuentes de datos oficiales y publicos que pueden integrarse: (1) El Sistema de Informacion y Reporte Empresarial (SIREM) de la Superintendencia de Sociedades, que contiene estados financieros NIIF de empresas colombianas (~9 GB de datos), y (2) La base de datos de empresas activas de la Camara de Comercio de Ibague, que contiene los NITs y clasificacion de empresas de la ciudad. Mediante un JOIN por NIT, es posible extraer especificamente los datos financieros de las PYMES ibaguenenses.

3.3 Justificacion tecnologica

Las tecnologias actuales de desarrollo web (React.js, Node.js, PostgreSQL) y las librerias de Machine Learning (Scikit-learn, XGBoost) permiten crear soluciones robustas a bajo costo. Ademas, el estudiante cuenta con recursos computacionales adecuados (Intel Core i5-10600KF, 16GB RAM, GPU RTX 2060) para procesar los volumenes de datos involucrados y entrenar modelos de aprendizaje automatico.

3.4 Justificacion academica

Este proyecto integra multiples areas de conocimiento de la Ingenieria de Sistemas: desarrollo web full-stack, bases de datos, procesamiento de datos, Machine Learning, y experiencia de usuario. Adicionalmente, incorpora conocimientos transversales de finanzas corporativas y normatividad contable (NIIF). La inclusion de un componente de inteligencia artificial diferencia este trabajo de proyectos tradicionales de desarrollo de software, aportando innovacion y rigor academico.

4. OBJETIVOS

4.1 Objetivo general

Disenar e implementar un sistema de informacion web que permita el analisis automatizado de estados financieros bajo normas NIIF y la evaluacion de riesgo mediante Machine Learning, generando reportes con indicadores, alertas y recomendaciones para apoyar la toma de decisiones en PYMES del municipio de Ibague, Tolima.

4.2 Objetivos especificos

OE1: Analizar los requerimientos funcionales y no funcionales del sistema mediante revision documental de la normatividad NIIF y entrevistas con empresarios de PYMES ibaguenenses.

OE2: Disenar la arquitectura del sistema aplicando patrones de diseno de software que garanticen escalabilidad, seguridad y mantenibilidad.

OE3: Integrar los datos del SIREM (Supersociedades) con la base de empresas de la Camara de Comercio de Ibague mediante un proceso ETL que permita identificar las PYMES de la ciudad.

OE4: Implementar el modulo de captura y procesamiento de estados financieros (Balance General, Estado de Resultados, Flujo de Efectivo) con calculo automatico de indicadores.

OE5: Desarrollar un modelo de Machine Learning para la clasificacion del nivel de riesgo financiero de las empresas, entrenado con datos historicos de PYMES colombianas.

OE6: Crear el modulo de visualizacion con dashboards interactivos, sistema de alertas semaforizadas y generacion de reportes ejecutivos en PDF.

OE7: Validar el sistema mediante pruebas funcionales, de rendimiento y de usabilidad con empresarios y contadores de PYMES de Ibague.

5. ALCANCE Y DELIMITACION

5.1 Delimitacion geografica

El proyecto se delimita al **municipio de Ibague, departamento del Tolima, Colombia**. Esta delimitacion se justifica por: (a) disponibilidad de datos especificos mediante el cruce SIREM - Camara de Comercio de Ibague, (b) acceso a usuarios para validacion y pruebas de usabilidad, (c) relevancia e impacto directo en la comunidad empresarial local, y (d) alcance manejable para un proyecto de pregrado.

5.2 Alcance funcional

| Modulo | Funcionalidades incluidas |
|---------------------|---|
| Autenticacion | Registro, login, recuperacion de contrasena, roles (empresario, contador, administrador) |
| Gestion de empresas | CRUD de empresas, asociacion de periodos contables, importacion de datos |
| Estados financieros | Captura/importacion de Balance General, Estado de Resultados y Flujo de Efectivo bajo estructura NIIF |
| Motor de calculo | Calculo automatico de 15+ indicadores financieros en 4 categorias (liquidez, rentabilidad, endeudamiento, eficiencia) |
| Machine Learning | Modelo de clasificacion de riesgo (bajo/medio/alto) con probabilidad porcentual |
| Visualizacion | Dashboard interactivo, graficos de tendencias, semaforos de alerta por indicador |
| Reportes | Generacion de informes PDF con analisis, recomendaciones y proyecciones |

5.3 Limitaciones y exclusiones

- No se incluye integracion automatica con software contable (Siigo, World Office, Helisa).
- No se desarrollara modulo de facturacion electronica ni conexion con la DIAN.
- El modelo de ML se entrenara con datos historicos; no realizara predicciones en tiempo real del mercado.
- Los datos de SIREM solo incluyen empresas vigiladas por Supersociedades; microempresas no obligadas a reportar no estaran disponibles.
- El sistema sera web responsive; no se desarrollara aplicacion movil nativa.
- Idioma unico: espanol.

6. MARCO TEORICO PRELIMINAR

6.1 Estados financieros bajo NIIF

Las Normas Internacionales de Informacion Financiera (NIIF) establecen los lineamientos globales para la preparacion y presentacion de estados financieros. En Colombia, las PYMES aplican la NIIF para PYMES (Grupo 2), segun lo establecido en el Decreto 3022 de 2013 y sus modificaciones. Los estados financieros basicos son:

| Estado Financiero | Proposito | Componentes clave |
|--------------------------------|--|------------------------------------|
| Estado de Situacion Financiera | Mostrar la situacion patrimonial a una fecha | Activos, Pasivos, Patrimonio |
| Estado de Resultado Integral | Mostrar la rentabilidad del periodo | Ingresos, Costos, Gastos, Utilidad |
| Estado de Flujo de Efectivo | Mostrar los movimientos de efectivo | Operacion, Inversion, Financiacion |

6.2 Indicadores financieros

| Categoria | Indicadores | Proposito |
|---------------|---|---|
| Liquidez | Razon corriente, Prueba acida, Capital de trabajo | Evaluar capacidad de pago a corto plazo |
| Rentabilidad | ROA, ROE, Margen bruto, Margen neto, Margen operacional | Medir capacidad de generar utilidades |
| Endeudamiento | Razon de deuda, Apalancamiento, Cobertura de intereses | Evaluar nivel de deuda y riesgo |
| Eficiencia | Rotacion de activos, Rotacion de inventarios, Rotacion de cartera | Medir uso eficiente de recursos |

6.3 Machine Learning para evaluacion de riesgo

El Machine Learning (aprendizaje automatico) permite construir modelos predictivos a partir de datos historicos. Para la evaluacion de riesgo financiero, se emplean modelos de clasificacion supervisada que, a partir de los indicadores financieros, predicen la probabilidad de que una empresa enfrente dificultades. Los algoritmos mas utilizados incluyen Regresion Logistica, Random Forest y XGBoost, este ultimo reconocido por su alto rendimiento en competencias de ciencia de datos.

7. FUENTES DE DATOS

7.1 SIREM - Superintendencia de Sociedades

El Sistema de Informacion y Reporte Empresarial (SIREM) contiene estados financieros de empresas colombianas bajo NIIF, publicados en Datos Abiertos Colombia:

| Dataset | Tamano | Contenido |
|--------------------------------|---------|--|
| Caratula NIIF | 2.07 GB | Metadatos de empresas y reportes |
| Estado de Situacion Financiera | 4.10 GB | Balance General (activos, pasivos, patrimonio) |
| Estado de Resultado Integral | 1.57 GB | Ingresos, costos, gastos, utilidades |
| Estado de Flujo de Efectivo | 1.43 GB | Flujos de operacion, inversion, financiacion |

7.2 Camara de Comercio de Ibague

Base de datos de empresas activas y matriculadas en la jurisdiccion de la Camara de Comercio de Ibague, disponible en Datos Abiertos Colombia. Contiene: NIT, razon social, actividad economica (CIU), tamano empresarial, municipio y datos de contacto. Esta base permite identificar especificamente las empresas de Ibague.

7.3 Estrategia de integracion (JOIN)

La integracion de ambas fuentes se realizara mediante un JOIN por el campo NIT:

1. Cargar base de Camara de Comercio de Ibague (NITs de empresas locales)
2. Filtrar: PUNTO_ENTRADA = 'NIIF Pymes' en datasets SIREM
3. JOIN: SIREM.NIT = CamaraComercio.NIT
4. Resultado: Estados financieros de PYMES especificamente de Ibague

8. ARQUITECTURA PROPUESTA

8.1 Stack tecnologico

| Capa | Tecnologia | Justificacion |
|------------------|---------------------------------|---|
| Frontend | React.js 18 + Tailwind CSS | Componentes reutilizables, responsive, moderno |
| Backend | Node.js + Express.js | JavaScript full-stack, alto rendimiento, API REST |
| Base de datos | PostgreSQL 15 | Robusta, gratuita, ideal para datos financieros |
| ORM | Prisma | Type-safe, migraciones automaticas |
| Autenticacion | JWT + bcrypt | Estandar de la industria, seguro |
| Machine Learning | Python + Scikit-learn + XGBoost | Librerias maduras, bien documentadas |
| Graficos | Recharts / Chart.js | Visualizaciones interactivas |
| Reportes PDF | jsPDF + html2canvas | Generacion client-side |
| Despliegue | Railway / Render / Vercel | Hosting economico, CI/CD integrado |

8.2 Diagrama de arquitectura de alto nivel

```
+-----+ +-----+ +-----+
| FRONTEND | | BACKEND | | BASE DATOS |
| (React.js) |<--->| (Node.js) |<--->| (PostgreSQL) |
| | | |
| - Dashboard | | - API REST | | - Empresas |
| - Formularios | | - Autenticacion | | - Estados Fin. |
| - Graficos | | - Logica negocio | | - Indicadores |
| - Reportes | | - ML Service | | - Usuarios |
+-----+ +-----+ +-----+
|
v
+-----+
| ML MODULE |
| (Python) |
| |
| - Modelo XGBoost |
| - API prediccion |
+-----+
```

9. METODOLOGIA

9.1 Tipo de investigacion

El proyecto se enmarca en una **investigacion aplicada de tipo desarrollo tecnologico**, con un enfoque mixto (cuantitativo para el analisis de datos y el modelo de ML, cualitativo para la evaluacion de usabilidad). Se seguira la metodologia agil **SCRUM** para el desarrollo del software.

9.2 Fases del proyecto

| Fase | Actividades | Entregables |
|-----------------------------|---|---|
| 1. Inicio (Sem 1-2) | <ul style="list-style-type: none">- Recopilacion de requerimientos- Descarga y exploracion de datos- Definicion de arquitectura | <ul style="list-style-type: none">- Documento de requerimientos- Dataset integrado- Documento de arquitectura |
| 2. Diseno (Sem 3-4) | <ul style="list-style-type: none">- Modelo de base de datos- Diagramas UML- Mockups de interfaces- Diseno del modelo ML | <ul style="list-style-type: none">- Modelo E-R- Diagramas de casos de uso- Prototipos de UI- Especificacion del modelo |
| 3. Desarrollo (Sem 5-12) | <ul style="list-style-type: none">- Implementacion por sprints- Desarrollo frontend/backend- Entrenamiento modelo ML- Integracion de componentes | <ul style="list-style-type: none">-Codigo fuente- Modelo ML entrenado- Sistema funcional |
| 4. Pruebas (Sem 13-14) | <ul style="list-style-type: none">- Pruebas unitarias- Pruebas de integracion- Pruebas de usabilidad- Validacion del modelo | <ul style="list-style-type: none">- Reportes de pruebas- Metricas del modelo- Resultados de usabilidad |
| 5. Cierre (Sem 15-16) | <ul style="list-style-type: none">- Despliegue en produccion- Documentacion final- Elaboracion del informe | <ul style="list-style-type: none">- Sistema desplegado- Manual de usuario- Documento de grado |

9.3 Metodologia para el modelo de Machine Learning

El desarrollo del modelo de ML seguira la metodologia CRISP-DM (Cross Industry Standard Process for Data Mining):

- **Comprension del negocio:** Definir que significa 'riesgo financiero' para una PYME.
- **Comprension de datos:** Explorar los datos SIREM, identificar variables relevantes.
- **Preparacion de datos:** Limpieza, transformacion, calculo de indicadores, creacion de etiquetas.
- **Modelado:** Entrenar y comparar modelos (Logistic Regression, Random Forest, XGBoost).
- **Evaluacion:** Medir precision, recall, F1-score, AUC-ROC; validacion cruzada.
- **Despliegue:** Integrar el modelo como microservicio en la arquitectura del sistema.

10. CRONOGRAMA

10.1 Cronograma detallado por sprints

| Sprint | Semana | Actividades |
|----------|--------|--|
| Sprint 0 | 1-2 | Configuracion del entorno, descarga de datos SIREM y Camara de Comercio, integracion (JOIN), exploracion inicial |
| Sprint 1 | 3-4 | Diseño de base de datos, diagramas UML, mockups de interfaces principales |
| Sprint 2 | 5-6 | Módulo de autenticación (registro, login, roles), estructura base del proyecto |
| Sprint 3 | 7-8 | Módulo de gestión de empresas y periodos, importación de datos |
| Sprint 4 | 9-10 | Captura de estados financieros (Balance, Resultados, Flujo Efectivo), validaciones |
| Sprint 5 | 11-12 | Motor de cálculo de indicadores financieros, desarrollo y entrenamiento del modelo ML |
| Sprint 6 | 13-14 | Dashboard interactivo, gráficos, sistema de alertas, integración del modelo ML |
| Sprint 7 | 15-16 | Generación de reportes PDF, pruebas finales, despliegue, documentación |

10.2 Diagrama de Gantt simplificado

[illegible]

11. PRESUPUESTO

| Rubro | Descripción | Cantidad | Valor Unit. | Total (COP) |
|-------------|---------------------------------------|-----------|---------------|------------------|
| Personal | Trabajo del estudiante (16 sem x 20h) | 320 horas | Aporte propio | \$0 |
| Equipos | Computador (ya disponible) | 1 | Disponible | \$0 |
| Datos | Descarga SIREM + Camara Comercio | - | Gratuito | \$0 |
| Hosting | Railway/Render (4 meses) | 4 meses | \$50,000/mes | \$200,000 |
| Dominio | Dominio .com.co (1 ano) | 1 | \$80,000 | \$80,000 |
| Software | VS Code, PostgreSQL, Python (OSS) | - | Gratuito | \$0 |
| Papeleria | Impresiones, empastado | - | - | \$150,000 |
| Transporte | Visitas validacion usuarios Ibague | 10 | \$15,000 | \$150,000 |
| Imprevistos | 10% del total | - | - | \$58,000 |
| | | | TOTAL | \$638,000 |

12. RESULTADOS ESPERADOS

- **Sistema web funcional:** Plataforma accesible via navegador que permita a PYMES de Ibague analizar sus estados financieros.
- **Modelo de ML entrenado:** Clasificador de riesgo financiero con precision $\geq 75\%$ (AUC-ROC).
- **Dashboard interactivo:** Visualizaciones graficas de indicadores con sistema de alertas semaforizadas.
- **Generador de reportes:** Informes PDF descargables con analisis, indicadores y recomendaciones.
- **Documentacion tecnica:** Manual de usuario, manual tecnico, documentacion de API.
- **Documento de grado:** Informe final con marco teorico, metodologia, desarrollo, resultados y conclusiones.
- **Articulo academico (opcional):** Paper para posible publicacion en revista indexada o congreso.

13. REFERENCIAS BIBLIOGRAFICAS PRELIMINARES

- [1] Camara de Comercio de Ibagué. (2024). Base de datos de empresas activas. Datos Abiertos Colombia. <https://www.datos.gov.co>
- [2] Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. KDD '16. <https://doi.org/10.1145/2939672.2939785>
- [3] Decreto 3022 de 2013. Por el cual se reglamenta la Ley 1314 de 2009 sobre el marco técnico normativo para los preparadores de información financiera. Presidencia de Colombia.
- [4] Gitman, L. J., & Zutter, C. J. (2016). Principios de administración financiera (14a ed.). Pearson Educación.
- [5] IASB. (2015). NIIF para las PYMES. Fundación IFRS. <https://www.ifrs.org/issued-standards/ifrs-for-smes/>
- [6] Laudon, K. C., & Laudon, J. P. (2020). Management Information Systems: Managing the Digital Firm (16th ed.). Pearson.
- [7] Ley 590 de 2000. Por la cual se dictan disposiciones para promover el desarrollo de la micro, pequeña y mediana empresa. Congreso de Colombia.
- [8] Ley 905 de 2004. Por medio de la cual se modifica la Ley 590 de 2000. Congreso de Colombia.
- [9] Ortiz Anaya, H. (2018). Análisis financiero aplicado bajo NIIF (16a ed.). Universidad Externado de Colombia.
- [10] Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. JMLR 12, 2825-2830.
- [11] Pressman, R. S., & Maxim, B. R. (2015). Software Engineering: A Practitioner's Approach (8th ed.). McGraw-Hill.
- [12] Superintendencia de Sociedades. (2024). SIREM - Sistema de Información y Reporte Empresarial. <https://www.supersociedades.gov.co>
- [13] Superintendencia de Sociedades. (2024). Estados Financieros NIIF. Datos Abiertos Colombia. <https://www.datos.gov.co>