**PROPUESTAS DE PROYECTO DE GRADO**

**Análisis de Datos de Precios de Alimentos en Bolivia**

**PROPUESTA 1: PREDICCIÓN DE INFLACIÓN DE ALIMENTOS POR REGIONES**

**Modelos: Regresión Lineal + K-fold Cross Validation**

**1. Introducción al Tema**

La seguridad alimentaria en Bolivia enfrenta desafíos significativos debido a la volatilidad de precios de productos básicos como papa, arroz, trigo y otros cereales. La inflación alimentaria afecta desproporcionalmente a los sectores más vulnerables de la población, especialmente en áreas rurales y periurbanas. El análisis predictivo de precios de alimentos utilizando técnicas de ciencia de datos puede proporcionar herramientas valiosas para la planificación de políticas públicas y la toma de decisiones estratégicas en seguridad alimentaria.

**2. Identificación del Problema**

Las autoridades gubernamentales y organizaciones de desarrollo carecen de herramientas predictivas confiables para anticipar aumentos significativos en los precios de alimentos básicos. Esta falta de capacidad predictiva resulta en respuestas reactivas tardías, afectando la estabilidad económica familiar y la seguridad alimentaria nacional. La ausencia de modelos cuantitativos que consideren variables geográficas, temporales y de mercado limita la efectividad de las intervenciones preventivas.

**3. Objetivo General**

Desarrollar un modelo predictivo de inflación de precios para productos alimentarios básicos en diferentes regiones de Bolivia, utilizando técnicas de regresión lineal y validación cruzada para mejorar la capacidad de planificación en seguridad alimentaria.

**4. Objetivos Específicos**

* Analizar patrones históricos de precios de alimentos básicos por departamento y mercado
* Identificar variables predictoras significativas incluyendo factores temporales, geográficos y de categoría de producto
* Implementar modelos de regresión lineal para predecir variaciones de precios a corto y mediano plazo
* Validar la precisión del modelo mediante técnicas de K-fold cross validation
* Desarrollar un sistema de alerta temprana para aumentos proyectados superiores al 20% mensual

**5. Límites y Alcances del Trabajo de Posgrado**

**Límites:**

* Se limitará al análisis de productos básicos incluidos en el dataset del WFP (cereales, tubérculos, legumbres)
* El período de análisis abarca desde 2008 hasta la fecha más reciente disponible
* No incluye factores externos como variables climáticas o políticas comerciales específicas
* Se enfoca únicamente en precios mayoristas documentados

**Alcances:**

* Cobertura de los 9 departamentos de Bolivia con datos disponibles
* Análisis de al menos 15 productos alimentarios básicos
* Desarrollo de modelos predictivos con horizonte de 1-6 meses
* Evaluación de precisión predictiva por región y tipo de producto

**6. Descripción del Trabajo de Posgrado**

El trabajo consistirá en cuatro fases principales: (1) Análisis exploratorio de datos para identificar patrones temporales y espaciales en los precios, incluyendo estacionalidad y tendencias regionales; (2) Ingeniería de características para crear variables predictoras como índices de volatilidad histórica, diferencias interregionales y factores estacionales; (3) Desarrollo y entrenamiento de modelos de regresión lineal múltiple con selección de características, implementando validación cruzada K-fold para optimizar hiperparámetros y evaluar robustez; (4) Validación y despliegue del modelo mediante pruebas con datos históricos reservados y desarrollo de un prototipo de sistema de alerta que genere reportes automáticos cuando se detecten proyecciones de inflación significativa.

**PROPUESTA 2: SISTEMA DE CLASIFICACIÓN DE MERCADOS POR VOLATILIDAD DE PRECIOS**

**Modelos: Linear Discriminant Analysis + Regresión Logística**

**1. Introducción al Tema**

La diversidad geográfica y económica de Bolivia genera heterogeneidad significativa en los mercados alimentarios locales. Algunos mercados presentan precios estables mientras otros experimentan alta volatilidad, afectando la planificación logística y las estrategias de intervención gubernamental. La clasificación sistemática de mercados según su comportamiento de precios permite optimizar recursos y diseñar políticas diferenciadas de seguridad alimentaria.

**2. Identificación del Problema**

Las instituciones responsables de la seguridad alimentaria aplican estrategias uniformes sin considerar las características específicas de volatilidad de cada mercado. Esta aproximación generalizada resulta en asignación ineficiente de recursos, intervenciones inadecuadas y falta de priorización en mercados críticos. La ausencia de una taxonomía científica de mercados según su estabilidad de precios impide el desarrollo de políticas públicas diferenciadas y efectivas.

**3. Objetivo General**

Desarrollar un sistema de clasificación automática de mercados alimentarios bolivianos según su nivel de volatilidad de precios, utilizando técnicas de análisis discriminante lineal y regresión logística para optimizar estrategias de intervención en seguridad alimentaria.

**4. Objetivos Específicos**

* Calcular métricas de volatilidad de precios para cada mercado incluido en el dataset
* Identificar características discriminantes de mercados (ubicación geográfica, diversidad de productos, patrones estacionales)
* Implementar modelos de Linear Discriminant Analysis para clasificación inicial de mercados
* Desarrollar modelos de regresión logística para refinamiento de la clasificación
* Crear un sistema de scoring para evaluar el riesgo de volatilidad de nuevos mercados

**5. Límites y Alcances del Trabajo de Posgrado**

**Límites:**

* Clasificación basada únicamente en datos de precios históricos disponibles
* No incluye factores socioeconómicos externos o características físicas de mercados
* Se limita a mercados con al menos 24 meses de datos continuos
* Clasificación en máximo 4 categorías de volatilidad

**Alcances:**

* Análisis de aproximadamente 50+ mercados con datos suficientes
* Desarrollo de taxonomía aplicable a mercados de toda Bolivia
* Sistema escalable para incorporar nuevos mercados
* Validación de clasificación con métricas de precisión superiores al 85%

**6. Descripción del Trabajo de Posgrado**

El desarrollo incluirá cinco etapas: (1) Cálculo de indicadores de volatilidad incluyendo desviación estándar, coeficiente de variación, rangos intercuartílicos y métricas de estacionalidad para cada mercado; (2) Análisis de componentes principales para reducir dimensionalidad y identificar factores latentes de volatilidad; (3) Implementación de Linear Discriminant Analysis para clasificación inicial, determinando fronteras de decisión entre categorías de volatilidad; (4) Refinamiento mediante regresión logística multinomial para mejorar precisión de clasificación; (5) Desarrollo de sistema de scoring que permita clasificar automáticamente nuevos mercados y generar alertas cuando un mercado cambie de categoría de riesgo.

**PROPUESTA 3: DETECTOR DE ANOMALÍAS EN PRECIOS - SISTEMA DE ALERTA TEMPRANA**

**Modelos: Naive Bayes + Greedy Search para optimización**

**1. Introducción al Tema**

Las crisis alimentarias frecuentemente se manifiestan a través de picos anómalos en los precios antes de convertirse en problemas de disponibilidad. La detección temprana de estas anomalías puede prevenir situaciones de inseguridad alimentaria y permitir intervenciones oportunas. Los sistemas de alerta temprana basados en análisis automatizado de patrones de precios representan una herramienta crucial para la gestión proactiva de la seguridad alimentaria nacional.

**2. Identificación del Problema**

Bolivia carece de un sistema automatizado de detección de anomalías en precios de alimentos que permita identificar situaciones de riesgo antes de que se conviertan en crisis alimentarias. Las alertas actuales dependen de reportes manuales tardíos, resultando en respuestas reactivas cuando los problemas ya han escalado. La falta de criterios objetivos y automatizados para definir qué constituye una anomalía de precios limita la efectividad de las intervenciones preventivas.

**3. Objetivo General**

Desarrollar un sistema automatizado de detección de anomalías en precios de alimentos para Bolivia, utilizando algoritmos de clasificación Naive Bayes y optimización Greedy Search para generar alertas tempranas de potenciales crisis alimentarias.

**4. Objetivos Específicos**

* Establecer líneas base de precios normales para cada producto y mercado usando análisis estadístico
* Implementar algoritmos Naive Bayes para clasificar precios como "normal", "alerta" o "crisis"
* Optimizar umbrales de detección mediante técnicas Greedy Search para minimizar falsos positivos
* Desarrollar sistema de scoring de riesgo agregado por región y producto
* Crear dashboard automatizado de alertas con actualización en tiempo real

**5. Límites y Alcances del Trabajo de Posgrado**

**Límites:**

* Detección basada únicamente en patrones históricos de precios
* No incorpora factores causales externos (clima, políticas, eventos sociales)
* Sistema reactivo a cambios de precios ya ocurridos
* Limitado a productos y mercados con datos históricos suficientes

**Alcances:**

* Monitoreo de todos los productos básicos en el dataset WFP
* Cobertura nacional con especificidad por departamento y mercado
* Sistema escalable para incorporar nuevas fuentes de datos
* Alertas diferenciadas por nivel de severidad y tipo de producto

**6. Descripción del Trabajo de Posgrado**

El proyecto se estructurará en seis módulos: (1) Análisis estadístico de distribuciones de precios para establecer rangos normales usando percentiles, medias móviles y análisis de estacionalidad; (2) Ingeniería de características para crear variables predictoras de anomalías incluyendo ratios de cambio, desviaciones de tendencia y comparaciones interregionales; (3) Implementación de clasificadores Naive Bayes multinomiales entrenados con datos históricos etiquetados; (4) Optimización de umbrales mediante algoritmos Greedy Search que maximicen sensibilidad mientras minimizan falsos positivos; (5) Desarrollo de sistema de agregación de alertas que combine señales de múltiples productos y mercados; (6) Construcción de dashboard web interactivo para visualización de alertas con capacidades de drill-down por región, producto y nivel de severidad.

**PROPUESTA 4: ANÁLISIS DE PATRONES DE CONSUMO REGIONAL**

**Modelos: Association Rules + Naive Bayes**

**1. Introducción al Tema**

Los patrones de consumo alimentario en Bolivia reflejan la diversidad cultural, geográfica y económica del país. La comprensión de estas correlaciones entre productos y regiones es fundamental para optimizar cadenas de suministro, planificar distribuciones de ayuda alimentaria y diseñar políticas de seguridad alimentaria culturalmente apropiadas. El análisis de reglas de asociación en precios revela comportamientos de consumo implícitos y dependencias entre productos.

**2. Identificación del Problema**

Las organizaciones humanitarias y gubernamentales carecen de conocimiento sistemático sobre los patrones de consumo regional y las interdependencias entre productos alimentarios en diferentes mercados bolivianos. Esta falta de comprensión resulta en estrategias de distribución ineficientes, programas de asistencia alimentaria que no reflejan preferencias culturales locales, y políticas de precios que no consideran efectos en cascada entre productos relacionados.

**3. Objetivo General**

Identificar y modelar patrones de consumo alimentario regional en Bolivia mediante análisis de reglas de asociación en comportamientos de precios y clasificación de regiones según perfiles de consumo utilizando algoritmos Naive Bayes.

**4. Objetivos Específicos**

* Aplicar algoritmos de reglas de asociación para identificar productos con comportamientos de precios correlacionados
* Mapear patrones de consumo predominantes por departamento y región cultural
* Clasificar mercados en perfiles de consumo usando algoritmos Naive Bayes
* Identificar productos sustitutos y complementarios mediante análisis de correlaciones de precios
* Desarrollar recomendaciones para optimización de cadenas de suministro regional

**5. Límites y Alcances del Trabajo de Posgrado**

**Límites:**

* Análisis basado en inferencias desde patrones de precios, no datos directos de consumo
* No incluye factores socioeconómicos o demográficos detallados
* Limitado a productos disponibles en el dataset del WFP
* Período de análisis restringido a datos disponibles (2008-presente)

**Alcances:**

* Cobertura de los 9 departamentos bolivianos
* Análisis de relaciones entre 15+ productos alimentarios básicos
* Identificación de al menos 5 perfiles regionales de consumo distintivos
* Generación de matriz de productos sustitutos/complementarios por región

**6. Descripción del Trabajo de Posgrado**

La investigación se desarrollará en cinco fases: (1) Preprocesamiento de datos para crear matrices de co-ocurrencia de variaciones de precios entre productos por mercado y período temporal; (2) Aplicación de algoritmos Apriori y FP-Growth para extraer reglas de asociación con métricas de soporte, confianza y lift significativas; (3) Análisis de correlaciones de precios para identificar productos sustitutos (correlación negativa) y complementarios (correlación positiva); (4) Implementación de clasificadores Naive Bayes para categorizar mercados en perfiles de consumo basados en composición de productos y patrones de correlación; (5) Validación de perfiles mediante análisis geográfico y cultural, desarrollando mapas de distribución de patrones de consumo y recomendaciones específicas para políticas alimentarias diferenciadas por región.

**PROPUESTA 5: MODELO INTEGRADO DE SEGURIDAD ALIMENTARIA**

**1. Introducción al Tema**

La seguridad alimentaria es un fenómeno multidimensional que requiere enfoques analíticos integrales para su comprensión y gestión efectiva. La combinación de múltiples técnicas de ciencia de datos permite crear sistemas comprehensivos que aborden diferentes aspectos del problema: desde la identificación de patrones hasta la predicción y clasificación de riesgos. Un modelo integrado proporciona una visión holística para la toma de decisiones estratégicas en política alimentaria nacional.

**2. Identificación del Problema**

Las instituciones responsables de seguridad alimentaria en Bolivia operan con sistemas fragmentados que abordan aspectos individuales del problema sin integración sistémica. Esta aproximación desarticulada genera visiones parciales, duplicación de esfuerzos, inconsistencias en la toma de decisiones y incapacidad para abordar la complejidad multifactorial de la inseguridad alimentaria. La ausencia de un sistema integrado de análisis limita la efectividad de las intervenciones y la optimización de recursos.

**3. Objetivo General**

Desarrollar un modelo integrado de análisis de seguridad alimentaria para Bolivia que combine múltiples técnicas de ciencia de datos para proporcionar una plataforma comprehensiva de monitoreo, predicción y clasificación de riesgos alimentarios.

**4. Objetivos Específicos**

* Integrar análisis de reglas de asociación, predicción de precios y clasificación de riesgos en un pipeline analítico unificado
* Desarrollar métricas compuestas de seguridad alimentaria que combinen múltiples indicadores
* Implementar sistema de validación cruzada integral para evaluar robustez del modelo combinado
* Crear dashboard interactivo que integre todas las funcionalidades analíticas
* Establecer protocolo de actualización y mantenimiento del sistema integrado

**5. Límites y Alcances del Trabajo de Posgrado**

**Límites:**

* Integración limitada a técnicas especificadas (association rules, Naive Bayes, regresiones, LDA, K-fold, Greedy search)
* Datos restringidos al dataset del WFP sin incorporación de fuentes externas
* No incluye factores macroeconómicos o políticos externos
* Sistema piloto sin implementación operacional completa

**Alcances:**

* Cobertura nacional completa con desagregación departamental y municipal
* Integración de todos los productos alimentarios disponibles en el dataset
* Sistema modular que permite activación independiente de componentes
* Validación integral con datos históricos de los últimos 5 años

**6. Descripción del Trabajo de Posgrado**

El desarrollo seguirá un enfoque de arquitectura modular en seis etapas: (1) **Módulo de Patrones:** Implementación de algoritmos de association rules para identificar interdependencias entre productos y regiones, creando una base de conocimiento de relaciones alimentarias; (2) **Módulo Predictivo:** Desarrollo de modelos de regresión lineal para proyección de precios integrados con el módulo de patrones para ajustar predicciones según correlaciones identificadas; (3) **Módulo de Clasificación:** Implementación de regresión logística y LDA para clasificar riesgo de inseguridad alimentaria por región, utilizando outputs de módulos anteriores como inputs; (4) **Sistema de Validación:** Implementación de K-fold cross validation para evaluar robustez de todo el pipeline integrado, con métricas específicas para cada módulo y métricas compuestas; (5) **Optimización Global:** Aplicación de greedy search para optimizar parámetros del sistema completo, balanceando precisión de diferentes módulos; (6) **Dashboard Integrado:** Desarrollo de interfaz web que integre todos los módulos con capacidades de visualización, alertas automáticas y generación de reportes ejecutivos para tomadores de decisiones.

**RECOMENDACIÓN FINAL**

Para una **primera presentación de proyecto de grado**, recomiendo la **Propuesta 3: Detector de Anomalías** por las siguientes razones:

✅ **Impacto práctico inmediato** - Sistema de alerta temprana con aplicación directa ✅ **Metodología clara** - Combinación equilibrada de técnicas especificadas  
✅ **Viabilidad técnica** - Alcances realistas para tiempo de posgrado ✅ **Relevancia social** - Aborda problema crítico de seguridad alimentaria ✅ **Escalabilidad** - Base para futuras extensiones del sistema

La propuesta permite demostrar dominio técnico mientras genera valor tangible para la sociedad boliviana.