

Docente Fred Torres Cruz
Autor: Jorge Luis Flores Turpo
Curso: Estadística Computacional
Video demostrativo: https://youtu.be/WD50y_d7nRs
Repositorio GitHub: https://github.com/Yorchisflrs/Analisis_discriminante

1. Actividad N° 016 – Análisis Discriminante PRO

1. Descripción general

El presente proyecto, titulado “**Análisis Discriminante PRO**”, consiste en el desarrollo de una aplicación web educativa e interactiva, capaz de ejecutar modelos estadísticos de clasificación supervisada como el Análisis Discriminante Lineal (LDA) y el Cuadrático (QDA). Fue construida con Python y Flask, integrando librerías como `scikit-learn`, `pandas`, `plotly` y `weasyprint`.

2. Objetivo

Facilitar el aprendizaje y aplicación práctica del análisis discriminante, mediante una interfaz amigable que permita:

- Cargar archivos `.csv` o `.xlsx`.
- Seleccionar variable objetivo.
- Aplicar modelos LDA o QDA con configuración mínima.
- Visualizar biplots, matriz de confusión y curva ROC.
- Generar e interpretar reportes en PDF con resultados estadísticos.

3. Fundamento teórico

El **Análisis Discriminante** es una técnica supervisada que asigna observaciones a grupos predefinidos. LDA asume homogeneidad en las matrices de covarianza y genera fronteras lineales, mientras que QDA permite matrices distintas por grupo, generando fronteras cuadráticas. Ambas técnicas buscan maximizar la separación entre clases.

4. Aplicación en el sistema

Los usuarios sin conocimientos técnicos pueden cargar sus datos y ejecutar modelos con pocos clics. La aplicación:

- Valida automáticamente las columnas y detecta la variable objetivo.
- Entrena modelos con validación cruzada.
- Presenta resultados gráficos y explicaciones pedagógicas.
- Permite exportar informes interpretativos en PDF.

5. Justificación del caso aplicado

El análisis se aplicó a un conjunto de datos educativos, donde se deseaba clasificar estudiantes según su rendimiento. Al aplicar LDA, se obtuvo una precisión del 82 %, y con QDA del 85 %, evidenciando mejor desempeño al considerar covarianzas distintas.

Interpretación: El modelo logró separar satisfactoriamente los grupos definidos (por ejemplo, “Aprobado” vs “Desaprobado”). El biplot evidenció la contribución de cada variable predictora. La matriz de confusión mostró una alta tasa de aciertos, y el AUC superior a 0.85 sugiere un modelo robusto.

6. Sostenibilidad y valor educativo

Este sistema no requiere instalación externa y respeta la privacidad: los datos se procesan localmente. Sirve como recurso autoformativo para estudiantes de estadística, ciencia de datos y carreras afines. La interfaz guiada incluye glosario, tooltips, y está diseñada para aprendizaje visual.

7. Conclusión

“Análisis Discriminante PRO” representa una propuesta sólida de integración entre teoría estadística y herramientas tecnológicas. Mejora la comprensión del análisis discriminante mediante una experiencia interactiva, accesible y educativa.

Palabras clave: Análisis discriminante, LDA, QDA, visualización, Flask, Estadística Computacional.