

Servicio web para la consulta de la tendencia de aprobación de materias

Ingeniería de Software 1

Diego Armando Rey¹

¹Facultad de Ingeniería
Maestra en Ciencias de la Información y las Comunicaciones
Universidad Distrital Francisco José de Caldas

Noviembre, 2017

1 Inicio

- Resumen
- Introducción

2 Desarrollo

- Análisis
- Implementación

1 Inicio

- Resumen
- Introducción

2 Desarrollo

- Análisis
- Implementación

- Diferentes herramientas estadísticas están siendo usadas en áreas sociales y comerciales. La precisión en los resultados ha llevado a su amplia aceptación e implementación. El área de la educación, aunque es el lugar donde emerge la teoría, no ha sido abordado a fondo mediante aplicaciones.
- Un estudio de aprobación de materias de estudiantes basado en observaciones e hipótesis de periodos previos se realiza en este documento. El método LDA(Linear Discriminant Analysis) ha sido escogido para establecer la regla de decisión dada la naturaleza de las observaciones. Posteriormente, se aplica el método a datos de cuatro materias diferentes con buenos resultados.

1 Inicio

- Resumen
- **Introducción**

2 Desarrollo

- Análisis
- Implementación

- Si un estudiante supiera que tan probable es aprobar una materia en el examen final antes de ingresar a presentarlo, tal vez cambiaría la cantidad de tiempo invertido en la preparación del examen. Si antes de inscribir una materia se conociera una regla que indique una nota mínima en cada evaluación para tener una *buena probabilidad* de aprobar la materia, tal vez cambiaría la cantidad de horas invertidas fuera del salón por el estudiante a lo largo del semestre.
- Tener una buena aproximación de la cantidad de estudiantes que van a aprobar o reprobado una materia ayudaría a dimensionar la carga docente del siguiente periodo académico o modificar la distribución de planta física. También, podrían programarse actividades correctivas como cursos vacacionales, tutorías intersemestrales o cursos cortos de refuerzo.

1 Inicio

- Resumen
- Introducción

2 Desarrollo

- Análisis
- Implementación

Con base en datos reales de notas de materias cursadas por estudiantes en varios semestres se realiz un modelamiento matemático para definir un clasificador por regla de decisión y se realizaron simulaciones en MATLAB®, para validar el anlisis, posteriormente en la etapa de implementacin se realizó la traducción a lenguaje R, el cual será el usado por el servidor RStudio®, para realizar el procesamiento de los algoritmos. Para determinar la regla de decisión se asume que los datos x_1 y x_2 se distribuyen normalmente (1)

$$\begin{aligned} \mathbf{X} &\sim \mathcal{N}(\mu_1, Q_T), \text{ para } H_1 \\ \mathbf{X} &\sim \mathcal{N}(\mu_2, Q_T), \text{ para } H_2 \end{aligned} \tag{1}$$

Se hizo uso de las técnica LDA¹. Buscando encontrar una buena regla de decisión, por tanto, se realizó la prueba con los datos de la *materia*₁ como se observa en la Fig. 1.

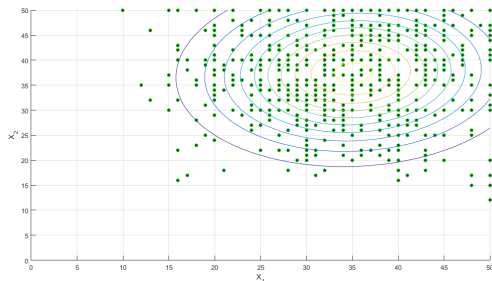


Figure: *Materia*₁: LDA, identificación de parámetros de gaussiana de H_1

¹Linear Discriminant Analysis

Como se observa en la Fig. 1 (note las escalas de los ejes), LDA es un método suficiente para establecer una buena regla de decisión dada la naturaleza del proceso (es matemáticamente imposible lograr aprobar la materia con ciertos valores de N_1 y N_2). LDA es el método utilizado para la prueba de hipótesis de cada materia.

Para obtener la regla de decisión mediante el método LDA, se obtienen los valores de μ_1 y Q_1 para la hipótesis H_1 y μ_2 y Q_2 para la hipótesis H_2 por materia a partir de los datos de los periodos previos.

Se establece una matriz de covarianza promedio en (2) para facilitar la obtención de la regla de decisión.

$$Q_T = \frac{Q_1 + Q_2}{2} \quad (2)$$

La recta que define la regla de decisión está dada por (3)

$$w^T x = c \quad (3)$$

Con w dado por (4) y c dado por (5)

$$w = (\mu_1 - \mu_2)^T Q_T^{-1} \quad (4)$$

$$c = \frac{1}{2} \left(\mu_1^T Q_T^{-1} \mu_1 - \mu_2^T Q_T^{-1} \mu_2 \right) \quad (5)$$

Una vez hallada la regla de decisión, se aplicó a cada uno de los estudiantes que en el periodo de prueba actual estaban cursando la materia. Se extraen los resultados para cada materia, generando los porcentajes de aprobación y reprobación en base a la prueba de hipótesis.

1 Inicio

- Resumen
- Introducción

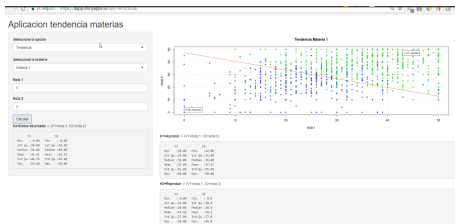
2 Desarrollo

- Análisis
- Implementación

Implementación

Se realizó la implementación y pruebas de la aplicación, utilizando tecnologías como Rstudio®, Microsoft Azure® y Shinyapps.io®. Las pruebas realizadas muestran el buen funcionamiento de la analítica de los datos consumidos desde el servicio web en Azure®. Figura 2.

Figure: Consumo desde ShinyApps a servicio web Azure



Conclusiones

- El método LDA fue suficiente para establecer una regla de decisión dadas las características de la información de periodos previos.
- Una fortaleza del estudio realizado se encuentra en su indiferencia al docente que imparte cada materia, sólo se basa en los resultados de las dos primeras evaluaciones en relación con su resultado final.
- Se incluyó el promedio ponderado del estudiante para mejorar el modelo, pero reveló que no tenía incidencia significativa en el resultado final de las materias.
- El servicio web de tendencia de materias puede proporcionar información pertinente a los estudiantes y directivos académicos para que tomen decisiones con respecto a su desempeño.

Conclusiones

- El tener un pronóstico acerca de un posible número de estudiantes reprobados por materia permite realizar la planeación de recursos en cursos de nivelación con por lo menos un mes de anterioridad a tener la nota definitiva de cada materia.
- El uso de tecnologías web permite ofrecer servicios de forma fácil y con disponibilidad de uso en distintas plataformas a un costo escalable según la demanda de consumo de información.
- Las plataformas de computación en la nube son una alternativa para el desarrollo de aplicaciones de forma ágil y con costos bajos para prueba de prototipos de software como solución de negocio en diferentes contextos de la industria de la información y las comunicaciones, con opción de escalamiento y mantenibilidad.

- Como trabajo futuro es posible desarrollar una regla de decisión más precisa mediante inferencia bayesiana para generar perfiles variables con menor error, e incluir más factores o variables que influyan en el resultado del proceso (aprobar o reprobar) y que permitan incluso predecir las notas de los estudiantes de forma aproximada al comportamiento de cada materia, buscando no tener un elevado costo computacional en la implementación de los algoritmos a utilizar.