

Examen Working Analyst

Módulo I: **clasificación**

¡Les felicitamos por haber terminado el primer módulo del curso de Machine Learning & AI for the Working Analyst!

La evaluación de cada uno de los módulos del curso está dividida en dos etapas, la primera es práctica y la segunda conceptual. Para aprobar este módulo es necesario aprobar ambas etapas y para aprobar el curso es necesario cumplir con la evaluación de por lo menos dos de los tres módulos.

Durante la semana de examen no habrán clases en vivo programadas, sin embargo los estudiantes pueden enviar sus dudas a los profesores por correo y en caso de ser necesario se agendarán reuniones con los profesores durante el horario de la clase. Es opcional realizar el examen en parejas o de manera individual y se deberá confirmar esta información el día martes de la semana del examen.

Examen conceptual _____

El examen conceptual lo presentarán en parejas o de manera individual el día viernes durante el horario de clase, la duración del examen es entre 15 y 20 minutos y será una sesión únicamente con el profesor Alfonso Ruiz.



- El examen conceptual NO es sobre matemáticas.
- El profesor presentará un caso de uso real que los estudiantes deberán desarrollar con algunos hints propuestos por el profesor.
- El caso de uso no necesariamente coincidirá con el del examen práctico.
- La evaluación estará sujeta al criterio del profesor.

Examen práctico

Se deberá entregar el día lunes antes de comenzar el siguiente módulo.

Por favor envíe el cuaderno de Colab con permisos para comentar y el resumen ejecutivo a los correo a [Ana Isabel](#) y a [Alfonso](#).

- Considere el conjunto de datos [FraudeCanastas.csv](#) disponible en el archivo zip correspondiente.
- Para conocer más información sobre el dataset original (no sobre el cual evaluaremos) pueden visitar [esta liga](#).
- Explore y prepare el dataset para construir un clasificador, eligiendo uno de los algoritmos vistos durante el módulo.
- Utilice un notebook de google colab donde además del código comente los hallazgos y pasos a seguir para construir un clasificador.
- Justifique la elección del algoritmo y los parámetros utilizados
- Evalúe el modelo
- Realice un resumen ejecutivo de no más de dos párrafos en un documento de texto.



Clasificación y ranking

(Venture Capital)

Les damos la bienvenida a nuestro examen conceptual sobre los problemas de clasificación. Pueden utilizar notas durante el examen.



Supongamos que un fondo de inversión llamado Bourbaki VC desea filtrar las solicitudes relacionadas con proyectos de inversión que podrían ser relevantes para su nivel de riesgo. Afortunadamente Bourbaki VC recibe muchas solicitudes por distintas vías, algunas de ellas son mensajería pero también podrían ser redes sociales o chats, esto ocasiona un *happy problem* pues la gestión de estas solicitudes es desgastante.

Dentro de Bourbaki VC trabaja una analista de datos que está familiarizada con Procesamiento del Lenguaje Natural y propuso construir un modelo de *machine learning* que filtre las solicitudes de manera automática y entregue a los directivos un reporte de las solicitudes más interesantes.

Palabras clave v.s. aprendizaje supervisado

El primer acercamiento es utilizar un filtro de búsqueda mediante algunas palabras clave, para encontrar estas palabras clave preguntó a los directivos cuáles son las palabras más llamativas entre las cuales se podrían destacar “patente”, “prototipo”, “estudio de mercado”, etc. Utilizando estas palabras clave podemos construir un reporte semanal con aquellos mensajes más interesantes y ahorraríamos mucho tiempo para Bourbaki VC.

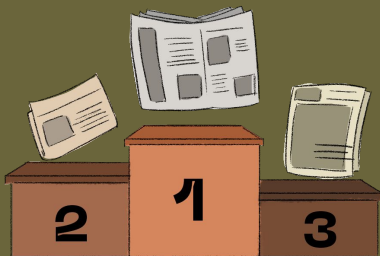
El segundo enfoque es construir un modelo de machine learning para clasificar de manera automática todas las solicitudes sin la necesidad de incluir una lista de las palabras clave. A los directores les gustaría utilizar Machine Learning sin embargo no tienen mucha experiencia en este tipo de temas.

PROBLEMA 1: ¿cuáles son las ventajas y dificultades del segundo enfoque? ¿existe alguna manera de fusionar ambos?

PROBLEMA 2: ¿cuál es el algoritmo que intuitivamente piensa que funcionaría mejor? ¿Por qué?

Ranking v.s. clasificación

Después de implementar el modelo de clasificación de machine learning los directivos están muy contentos porque los reportes que se han recibido hasta el momento son bastante interesantes e inclusive han invertido en algunos de ellos, de hecho han decidido construir un mayor grupo de ciencia de datos dentro de Bourbaki VC. Durante una de las reuniones con el equipo de Ciencia de Datos uno de los directivos pregunta si además de entregar la lista de las solicitudes interesantes, podrían hacer esta lista ordenada, es decir que las más relevantes para Bourbaki VC fueran al inicio del reporte, es decir que están rankeadas.



PROBLEMA 3: ¿Si el reporte utiliza un modelo de máquinas de soporte vectorial es posible construir un ranking que sustituya a la clasificación binaria? ¿Cuál modelo de clasificación podría construir este ranking?

Si quedaron algunas dudas sobre las preguntas durante el examen no olvides escribirle al Profesor Alfonso.

Quienes deseen continuar con su aprendizaje al terminar el curso les recomendamos inscribirse a alguna de las continuaciones^o.

Por favor acercarse a los profesores para preguntarles cuál de ellos es ideal para su perfil

1. Aplicaciones Financieras de ML & AI [Temario.](#)
2. Especialización en Deep Learning. [Temario.](#)
3. Track de Ciencia de Datos. [Temario.](#)

✉ info@colegio-bourbaki.com

☎ +52 999 267 7399

Colegio de Matemáticas Bourbaki

