

## Examen final

Érase una vez, en una tierra lejana, vivía una hermosa princesa que había caído gravemente enferma. El rey y la reina intentaron todo lo que estuvo a su alcance para curarla, pero ninguno de los remedios que probaron pareció funcionar.

Un grupo de aldeanos locales se acercó al rey y la reina y les hablaron de un conjunto de ingredientes mágicos que se decía que tenían el poder de curar cualquier dolencia. Sin embargo, los aldeanos advirtieron que estos ingredientes podrían ser volátiles en sus efectos, además, debido a las sequías recientes, solo algunos de los ingredientes pueden estar disponibles en un momento dado, y solo un alquimista experto podría determinar si una combinación específica de estos ingredientes particularmente volátiles y escasos curaría a la princesa.

El rey y la reina estaban desesperados por salvar a su hija, por lo que se dispusieron a buscar al mejor alquimista de la tierra. Buscaron por todas partes y finalmente encontraron a un alquimista que tenía la reputación de ser un maestro de un nuevo arte mágico conocido como "Machine Learning".

El alquimista primero examinó las afirmaciones de los aldeanos y las cantidades que habían tomado de cada ingrediente, junto con si había llevado o no a una cura. El alquimista sabía que esta era su única oportunidad de curar a la princesa, y tenía que hacerlo bien.



[Esta foto](#) de Autor desconocido está bajo licencia [CC BY-NC-ND](#)

Los ingredientes (características que usaremos para predecir) son los siguientes:

- Pluma de fénix
- Cuerno de unicornio
- Sangre de Dragon
- Lágrimas de sirena
- Polvo de hadas
- Dedos de duende
- Brebaje de bruja
- Garra de grifo
- Pelo de trol
- Tinta de kraken
- Cuerno de Minotauro

- Escala de basilisco
- Colmillo de quimera

Lo que nos permitirá determinar la variable objetivo que es:

- Curado (si la combinación de ingredientes y cantidad conduce a una cura)

Objetivo:

- Determinar que combinación de ingredientes nos ayudara a tener una mejor probabilidad de lograr el objetivo, curar a la princesa. Esto lo haremos creando un modelo con los ingredientes que creamos mas adecuados para tal fin. Recuerde que los ingredientes son escasos, por lo que determinar los que nos brinden mayor probabilidad de poderla curarla es clave.

## Tema I – 50 puntos

**Utilizando R, cargue el dataset y observe los datos antes de responder. (sugerencia: verifique los datos con la funcion summary)**

1. Este es un problema de clasificacion binaria o un problema de clasificacion multiclase? Cual es la diferencia entre ambos. **5 puntos**
2. En sus palabras porque es importante y primordial efectuar el proceso del feature engineering, para un problema de clasificacion? **5 puntos**
3. En un dataset con tantas características (features) diferentes, como puedo elegir de una manera objetiva (numerica) que features probar en el modelo y cuales no. **Explique su razonamiento. 5 puntos**
4. Como puede prevenir el overfitting o el underfitting en este caso? **15 puntos**
5. A lo largo de la ultima parte del curso, se expusieron varias metricas para medir el éxito del modelo de clasificacion (RMSE, Accuracy, precision, recall, f1 score), y teniendo en mente que el objetivo es intentar encontrar la combinacion de ingredientes que nos brinden la mejor probabilidad de encontrar una combinacion de ingredientes que salven a la princesa, cual de estas cuatro metricas seria la mas adecuada para poder medir el modelo. **EXPLIQUE SU RAZONAMIENTO. 20 puntos**

## Tema II - 50 puntos

Teniendo en cuenta el tiempo, desarrolle al menos tres experimentos de su modelo, aplicando lo visto en clase y tomando en cuenta lo consultado en el Tema I. Deje constancia de cada uno de los tres experimentos en su notebook de R, mejor si lo deja bien documentado.

En base a los experimentos ejecutados y basandose en la metrica elegida en la pregunta 5 del tema I, efectue la recomendación del modelo que mejor podría ayudarnos a predecir un éxito para poder curar a la princesa.

Nota: Recuerde somos ingenieros, por lo que debemos aplicar algunas de las mejores practicas vistas en clase.

Entregables:

- Notebook en R de su código y los experimentos y demás procedimientos efectuados.
- PDF con sus respuestas del tema I y las ejecuciones efectuadas del tema II.