**Data Scientist Qroma – Evaluación**

**Instrucciones generales**

* El examen tendrá una duración de 1 hora y media.
* Puedes realizar preguntas durante todo el examen.
* Envía este archivo con tus soluciones en el formato nombre\_apellido.docx al correo anapaula@brein.pe con el asunto solucion\_examen.

**Preguntas teóricas**

(1 punto cada una)

1. ¿Qué métricas conoces para evaluar modelos de clasificación? ¿En qué se diferencian y en qué casos de uso (ejemplos) elegirías cuál usar?

2. ¿Qué métodos de regularización conoces? Explica a detalle cómo funciona la regularización para un algoritmo en particular.

3. ¿Cuál es la diferencia entre bagging y boosting? ¿Qué algoritmos conoces que empleen estos métodos?

4. ¿Qué métodos de detección y tratamiento de outliers conoces?

5. ¿Cuál es la diferencia entre Ordinal Least Squares y Gradient Descent en el caso de la regresión lineal?

6. ¿Cómo se interpretan los coeficientes de una regresión logística?

7. Has realizado un modelo de regresión y deseas hacer inferencia sobre los coeficientes, ¿qué supuestos debes verificar?

8. ¿Qué métodos de correlación entre variables conoces? ¿En qué casos no utilizarías la correlación de Pearson?

9. Explica el principio de aleatorización en un diseño de experimentos y por qué es importante emplearlo.

10. Explica qué es el sesgo en el muestreo (sampling bias) y un ejemplo en donde los resultados son engañosos.

**Preguntas prácticas**

(5 puntos cada una)

1. Cuentas con una base de datos de hace 5 años del detalle de las ventas diarias de todos los productos de Qroma. De estos, aproximadamente 10 000 son clasificados como “Made to Stock”. Necesitas predecir la demanda mensual en galones de cada uno de estos productos en el siguiente mes para poder planificar la producción. Explica como plantearías este problema como uno analítico.

- ¿Qué tipo de modelo es? ¿Qué técnicas utilizarías?

- ¿Cómo evaluarías los resultados de tu modelo? ¿Qué métricas utilizarías? ¿Cómo repartirías la data?

- ¿Qué variables podrías crear para cada producto?

- ¿Cómo implementarías los resultados de tu modelo?

2. Cuentas con algunos indicadores del año pasado de los clientes de ferreterías de Qroma en el archivo “Data\_Clientes.csv”. El gerente de este canal desea entender (de manera descriptiva) cómo son sus clientes. Arma una pequeña presentación, infografía o reporte con los indicadores, estadísticas, cuadros o gráficos que creas conveniente. Eres libre de usar cualquier programa. **No te olvides de adjuntar cualquier archivo adicional con tu nombre\_apellido en el correo.**