

# Data Scientist

En *DataCompany* queremos hacer crecer nuestro equipo con personas que tengan ganas de desarrollar software basado en datos con dos objetivos: ayudar en la transformación a nuestros clientes y disfrutar del proceso de crear valor a través de la tecnología.

Si quieres ser parte de un equipo que te ofrecerá retos tecnológicos y que te exigirá una mejora y aprendizaje continuo, si no tienes miedo a equivocarte y eres capaz de ser proactivo y tirar del carro cuando te toque, es posible que *DataCompany* sea el sitio donde deberías trabajar. ¿Te interesa? Echa un vistazo a nuestras vacantes en el equipo de Data Science.

## ¿Sabes lo que vas a hacer además?

El objetivo del proyecto junto con el del equipo es realizar el desarrollo, entrenamiento y productivización de modelos para dar respuesta a necesidades de negocio del cliente. Todos los casos de uso que se llevarán a cabo estarán desarrollados en Cloud (AWS).

## Requisitos:

- Mínimo 1 años de experiencia en en el desarrollo de piezas de Machine Learning así como en su implementación en entornos de nube pública
- Programación en Python y/o R
- Manejo de librerías de análisis, modelado y visualización de datos.
- Conocimientos y experiencia trabajando con deep learning, especialmente NLP
- Fuerte base en estadística.
- Conocimientos en Git.
- Capacidad de síntesis y habilidades comunicativas en presentaciones.

## También es deseable:

- Experiencia en despliegue de modelos en producción
- Inglés nivel B2-C1.
- Experiencia y entusiasmo por trabajar con metodologías ágiles (Scrum, Kanban...).

## ¿Qué te ofrecemos en DataCompany?

- Un buen ambiente de trabajo junto a compañeros que son unos cracks en sus tecnologías y de los que aprenderás cada día.
- Tendrás un presupuesto personal anual de 2.000 euros para que lo dediques a la formación elegida por ti, adicional a cubrir el coste de tus certificaciones.
- Horario flexible, así como teletrabajo y flexibilidad en el almuerzo, para que tu vida personal y profesional encajen sin problema.
- Jornada intensiva los viernes todo el año, y los meses de julio y agosto.
- Retribución flexible para tarjeta transporte, seguro médico para familiares y cheque guardería.
- Tu evaluación y subida salarial la realizarán tus compañeros, nada de jefes porque aquí no encontrarás ninguno.

## **PRUEBA TÉCNICA**

Para evaluar tus conocimientos técnicos, en esta fase del proceso de selección tendrás que realizar un ejercicio práctico del que dispondrás hasta el viernes 10 donde se dará lugar a la entrevista individual.

### **Caso de uso**

Tenemos el caso de uso de un cliente (@TheBridge\_Tech) que desea desarrollar una monitorización de las redes sociales para medir el impacto de su marca y acciones comerciales.

Tendrás que realizar las siguientes tareas:

1. **Recopilar los tweets** donde se mencione la cuenta de @TheBridge\_Tech desde el día que comenzaste el bootcamp (21 Noviembre) hasta el último día de clase (7 Marzo). Se recomienda utilizar la [API v2 de Twitter](#). Se deberá recoger:
  - a. Id del mensaje
  - b. Cuerpo del texto del mensaje
  - c. Fecha del tweet
  - d. Id del autor
  - e. Nombre del autor
  - f. Nombre de usuario del autor
  - g. Métricas públicas del tweet (retweet, reply, like, quote)
2. Almacenarlos en una **base de datos SQL** desplegada en AWS en 2 tablas normalizadas: (tweets, usuarios)
3. Realizar un pequeño **análisis** donde se respondan a las siguientes preguntas de negocio:
  - a. ¿Cuál es el tweet con mayor repercusión social?
  - b. ¿Cuál es el usuario que más menciona a la escuela?
  - c. ¿En qué mes se concentra el mayor número de tweets?
  - d. ¿Qué palabras son más frecuentes?
  - e. ¿Qué tipo de correlación matemática encuentras entre las métricas públicas?
  - f. ¿Has sacado alguna conclusión extra en tu análisis?
4. Utiliza el **modelo pre entrenado** que se te ha facilitado de análisis de sentimiento para determinar el sentimiento de los tweets. Preguntas:
  - a. ¿Cuáles son las predicciones? Interpreta los resultados.
  - b. ¿Qué variables son las más importantes del modelo?
  - c. ¿Cómo podrías mejorar el modelo?
  - d. ¿Qué otras oportunidades se te ocurren donde se podrían aplicar otros modelos de ML?
5. Realizar una **presentación** de 5 minutos máximo, donde presentes tu solución y la defiendas en una entrevista personal.

Las herramientas o recursos que necesites para el desarrollo de la prueba técnica quedan a tu elección. Deberás subir tanto tu código como la presentación a un repositorio de Git.

**¡Mucha suerte!**