|  |
| --- |
| **ACTIVIDAD 2.3 – Ejercicio Teórico**  **Anteproyecto de minería de datos web** |
| Imagina que decides montar un sistema de minería de datos web basado en un motor predictivo de Machine Learning.  Este es un ejercicio teórico. Explica en una **extensión máxima de 1 página** (aproximado) tu idea de proyecto, intentando contestar en la medida de lo posible a cuantas puedas de las siguientes preguntas. No es necesario que escribas nada de código, pero si lo ves necesario puedes comentar algunas de las tecnologías que usarías también. |
| Nombre Alumno:  Otros Alumnos Grupo (si hay): |
| * ¿Cuál es el objetivo de tu proyecto? * ¿Qué tipo de minería web usarás: de contenido, de utilización, de estructura? * ¿Cuáles serían tus fuentes de datos? ¿Qué web o conjunto de webs usarías para extraer datos?   + ¿Usarías una web de un dominio propio o de tu empresa, o usarías webs externas?   + ¿Qué tipo de datos quieres minar: numéricos, textos, imágenes?   + Señala la web concreta y pon ejemplos concretos de datos en esa web querrías minar   + ¿Prevés algún problema en la adquisición de los datos? * ¿Con qué frecuencia tendrías que hacer el minado de datos?   + ¿Qué cantidad de datos esperas manejar para crear tu dataset de entrenamiento?   + ¿Y una vez esté el sistema en funcionamiento? * ¿Qué tipo de sistema predictivo (de Machine Learning) propones?   + ¿Qué cosa te interesa predecir?   + ¿Vas a usar sistemas de clustering, clasificación, regresión, análisis de textos?   + ¿Cuáles crees que son las métricas adecuadas para validar tu sistema de minería web? * Más allá de este ejercicio teórico, ¿cómo de viable crees que sería tu propuesta de sistema en la vida real? |
| **Propuesta de Projecto**  A continuación, describe tu proyecto: |