Brainstorm

Posibles ideas preseleccionadas:

**Cartas Pokemon**. Predictor de precio de cartas Pokemon al igual que Brickeconomy.

**Predictor de precios por barrio.** Un predictor del precio de la vivienda por barrios a partir de datos públicos disponibles (delincuencia, acceso a servicios públicos, etcétera)

**Predictor cotizaciones bolsa a partir de comentarios en redes sociales.** A partir del seguimiento de Tweets de CEOs creación de un algoritmo de compraventa de acciones.

**Predictor de impagos.** Indicador de cómo de cerca está la empresa de un impago a partir de fuentes públicas.

**Servicio de scoring de “fiabilidad” de una empresa**, A partir de un agregador de opiniones, comentarios en redes sociales, indicadores de satisfacción de usuarios, etcétera

La opción seleccionada es la última porque se ha visto más factible para el nivel del promotor.

Diseño del DAaaS

### Definición la estrategia del DAaaS

***“Truster.com”.*** *La plataforma permite, por suscripción, acceder a la fiabilidad y seriedad de cualquier empresa del registro mercantil de España como potencial socio comercial, estratégico o financiero. Para ello, combina diversas fuentes publicas y privadas disponibles para ofrecer un “scoring” de seriedad de la empresa.*

*La funcionalidad comercial de la solución ofrecida es múltiple. Los negocios se basan en la confianza mútua. El modelo pretende combinar indicadores cuantitativos y cualitativos para predecir cómo de fiable es la contrapartida. Puede ser útil tanto para hacer un filtro previo de (1) potenciales proveedores, (2) clientes que pagan en diferido, (3) socios comerciales.*

### Arquitectura DAaaS

*Los componentes utilizados en la construcción serán los siguientes.*

***Fuentes de inputs:***

1. *Indicadores objetivos de impago:*
   1. *Registro de Aceptaciones Impagadas (gestionado por Experian)*
   2. *Experian Bureau Empresarial*
   3. *Registro impagados judiciales*
   4. *Multas y sanciones administrativas en el Boletín Oficial del Estado*
2. *Indicadores de sentimientos:*
   1. *Linkedin corporativo y de los empleados*
   2. *Twitter corporativo y de los empleados. Respuestas recibidas.*
   3. *Noticias positivas o negativas en prensa online.*
   4. *Webs sobre atractivo para trabajar (glassdoor y go work)*
   5. *Webs sobre opiniones de clientes de empresa (Trustpilot y OpinionesEspaña)*
3. *Seriedad de la dirección:*
   1. *Multas y sanciones de administradores en el BOE*
   2. *Noticias en prensa online*
4. *Estado financiero de la empresa:*
   1. *Acceso vía SABI a las cuentas anuales de la empresa registradas en el registro mercantil*

***Herramientas:***

1. *Scrapping: conexión vía API (1a, 1b, 1c, 2a, 2b )*
2. *Crawling: Utilización de crawlers (Scrappy o Beautiful Soup) (1d, 2c, 2d, 2e, 3a, 3b)*
3. *Google Cloud Storage: para el almacenamiento de datos*
4. *Hadoop para el procesamiento de datos*
5. *Elastic search para filtrar los indicadores de sentimiento*
6. *Cloud functions para almacenar los algoritmos de screening*
7. *Tableau para la visualización de los datos*
8. *Wordpress, D3, programación en HTML para el diseño de la interfaz de usuario*

### DAaaS Operating Model Design and Rollout

*Personalizar los modelos operativos DAaaS para cumplir con los procesos, la estructura organizacional, las reglas y el gobierno de los clientes individuales. Realizar seguimiento de consumo y mecanismos de informe.*

Pasos esenciales:

1. Crear un Google Cloud Project y un bucket de Cloud Storage.
2. Guardar en Google functions los algoritmos necesarios:
   1. las búsquedas mediante crawling *(1d, 2c, 2d, 2e, 3a, 3b)*
   2. Los API rest necesarios *(1a, 1b, 1c, 2a, 2b )* según especificaciones de cada proveedor de información.
   3. La medición de sentimientos a través de elastic search sobre los inputs sobre la empresa *(2a, 2b, 2c, 2d, 2e)*
   4. Alertas de reputación para los administradores a través de elastic search sobre inputs de los administradores *(3b)*
   5. La puntuación de la empresa según datos *(1a, 1b, 1c, 1d)*
3. Almacenar toda la información del ecosistema empresarial español en Google Cloud Storage.
4. Depurar los datos almacenados a través de un proceso ELT.
5. Ejecutar mediante Hadoop (se asume necesidad de alta capacidad de computación):
   1. Los algoritmos de Elastic Search.
   2. Los algoritmos de screening/puntuación de las empresas para cada una de las categorías.
6. Almacenar resultados en Google SQL
7. Lanzamiento de queries a través de SQL a solicitud del cliente.

Link a Diagrama:

<https://drive.google.com/file/d/1xzOs3R2ORzVSl16nmnrh8_Qtzmxwi2T8/view?usp=sharing>

