

Una compañía de intermediación de seguros gestiona una cartera de 300.000 clientes. Para cada cliente almacena información sobre sus datos de contacto y las pólizas que tiene contratadas. Sobre estos datos, la dirección quiere tener cuadros de mando en los que poder obtener información sobre evolución de las pólizas contratadas, el desempeño de cada sucursal, etc. Además, les gustaría tener modelos predictivos que les permitan adelantarse a la demanda o prever clientes que podrían darse de baja. ¿Tendría algún beneficio desplegar una plataforma Hadoop en la compañía? En caso de ser beneficioso, ¿qué casos de uso por ejemplo podrían implementarse que ahora no se implementan?

Implementar hadoop sería muy beneficioso para la empresa por varios motivos:

- Gestión eficiente y sencilla del gran volumen de datos de la empresa.
- Bajo coste, debido a que se trata de una aplicación open source que no requiere un dispositivo de alto rendimiento.

Además, con esta aplicación se cumplirían todas las necesidades que busca la dirección de la empresa:

- Cuadro de mandos dinámicos que integran los datos a tiempo real y permiten una visualización sencilla de los datos.
- Incluye un modelo predictivo capaz de identificar clientes con alta probabilidad de cancelar sus pólizas mediante algoritmos. También podría prever las pólizas que serán más demandadas por regiones.

Por último, también se podrían implementar varios casos de uso que no estarían contemplados en la actualidad:

- Implementación de una nueva gestión tanto de los datos estructurados (cartera de clientes) como de datos no estructurados (mails, notas...).
- Integración de datos de fuentes externas, como redes sociales o comportamiento web.
- Agrupación de clientes por valor de datos para realizar campañas de marketing más efectivas.
- Creación de patrones capaz de detectar intentos de fraude.

En resumen, la implementación de hadoop sería muy beneficiosa tanto como para cumplir los objetivos de la dirección como para implementar varias mejoras que mejorarían la gestión, velocidad y comprensibilidad de los datos.