

Predicción de incumplimiento de préstamos, deep learning y data analytic.

Estructura del documento:

- Introducción
- Elección del conjunto de datos
- Enfoque deep learning
- Enfoque data analytic
- Conclusiones y aprendizajes

Introducción:

En el contexto de la gestión de riesgos financieros, la predicción de incumplimiento de préstamos se erige como un componente esencial para las instituciones bancarias. La complejidad de los datos financieros modernos demandan enfoques avanzados, y las técnicas de Deep Learning pueden ser muy prácticas.

La aplicación de Deep Learning, potencia la capacidad predictiva al permitir la identificación de patrones y correlaciones sutiles en datos financieros. Este enfoque ofrece una mejora significativa en la precisión de las predicciones, proporcionando una base sólida para la gestión eficaz de riesgos y decisiones informadas.

La versatilidad de esta combinación es esencial en un entorno financiero dinámico. La capacidad de los modelos resultantes para aprender y adaptarse a cambios en los comportamientos financieros asegura su relevancia continua ante las evoluciones económicas, posicionando a las instituciones para enfrentar los desafíos futuros con solidez.

Elección del conjunto de datos:

La elección del dataset de Lending Club para mi investigación en préstamos se fundamenta en diversas razones:

- **Volumen Significativo de Datos:**

El extenso dataset de Lending Club proporciona una abundancia de información, permitiéndome realizar análisis detallados y construir modelos sólidos.

- **Relevancia Financiera Auténtica:**

Lending Club, como una plataforma de préstamos entre pares establecida, ofrece datos que reflejan transacciones reales en un entorno financiero.

- **Diversidad de Información:**

La variada gama de datos en el dataset, incluyendo tasas de interés, historial crediticio, propósito de los préstamos e ingresos, brinda la oportunidad de explorar diversos factores que podrían influir en los incumplimientos.

- **Accesibilidad Pública:**

La disponibilidad pública de los datos de Lending Club facilita el uso del dataset.

- **Desafío Relevante en Finanzas:**

Abordar la predicción de incumplimientos de préstamos representa un desafío significativo y real en el sector financiero, brindándome la oportunidad de aportar soluciones prácticas.

Enfoque deep learning:

Metodología del enfoque deep learning:

Se realizará una transformación de características para convertir valores categóricos en valores numéricos discretos. Esto es esencial para permitir que el modelo de Deep Learning trabaje con todas las características de manera efectiva.

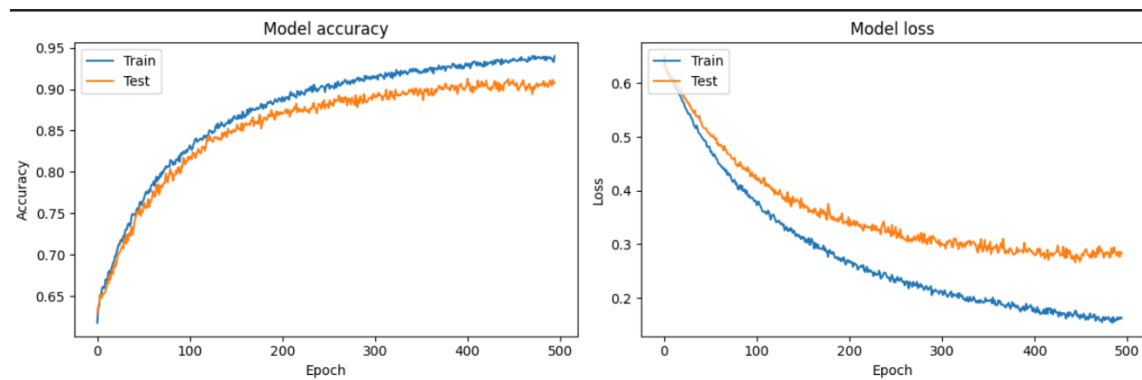
Se llevará a cabo un análisis exploratorio de datos para observar diferentes factores dentro del conjunto de datos. Esto incluirá la visualización de distribuciones, estadísticas descriptivas y la identificación de posibles patrones o tendencias.

Se realizarán verificaciones las correlaciones entre las diferentes características. Aquellas con una fuerte correlación se identificarán y se considerarán su eliminación para reducir el número de características, conservando solo aquellas más relevantes para el análisis.

Se construirá un modelo de Deep Learning utilizando la biblioteca TensorFlow. Este paso se realizará después de aplicar la transformación de características y realizar el análisis de los datos.

En resumen, el objetivo es desarrollar un modelo capaz de predecir la probabilidad de incumplimiento de préstamos futuros.

Resultados del enfoque deep learning:

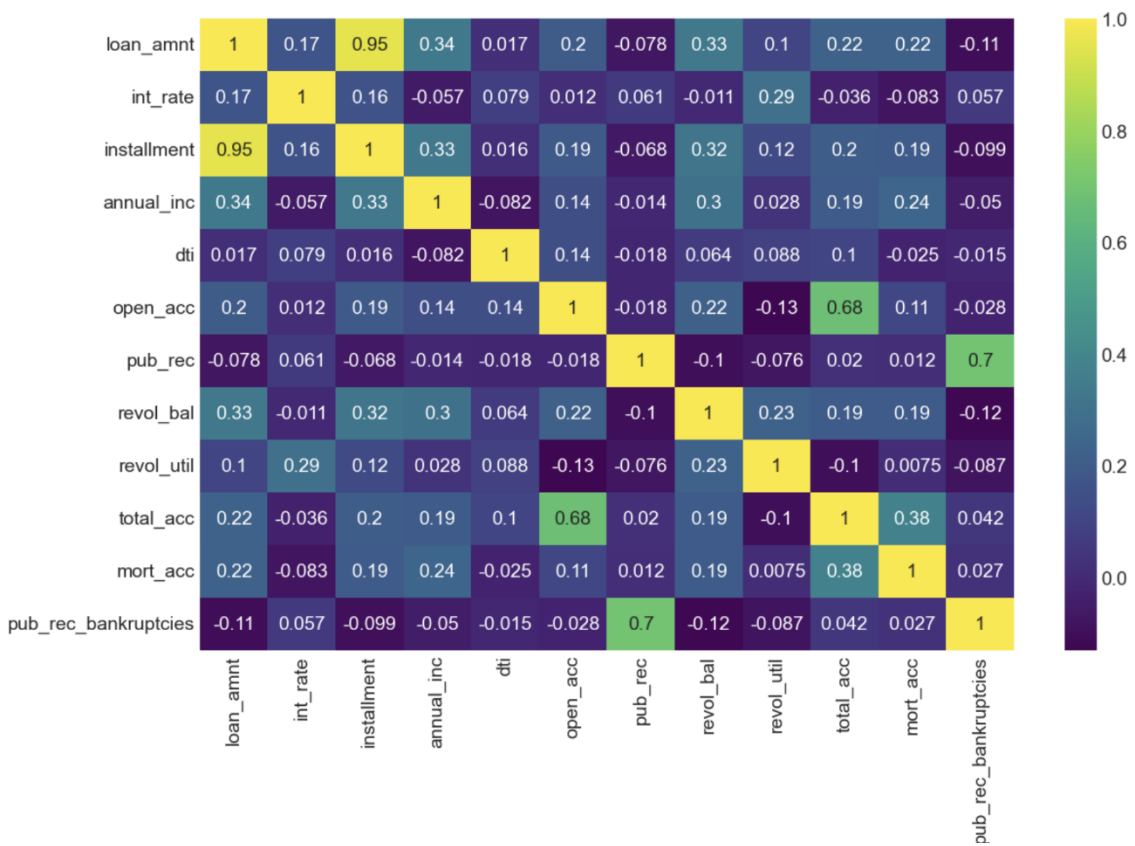


Enfoque de análisis de datos (data analytics):

Metodología del data analytics:

Se identificarán variables clave que podrán influir en la probabilidad de impago, como historial crediticio, ingresos, deudas, historial laboral, entre otros.

Se examinarán y visualizarán datos para comprender la distribución de variables, además de detectar patrones y explorar posibles correlaciones.



Conclusiones:

- **Importancia Estratégica:**

La predicción de incumplimiento de préstamos emerge como un componente estratégico en la gestión de riesgos financieros, ofreciendo a las instituciones bancarias una base sólida para decisiones informadas y mitigación de pérdidas.

- **Elección del Conjunto de Datos:**

La selección del dataset de Lending Club se fundamenta en su volumen significativo, relevancia financiera, diversidad de información y accesibilidad pública, consolidándolo como un recurso valioso para abordar la predicción de incumplimientos.

- **Logros en Deep Learning:**

La transformación de características y el análisis exploratorio de datos en el enfoque de Deep Learning posibilitan la construcción de un modelo robusto, hábil para anticipar la probabilidad de incumplimiento. La implementación eficaz se facilita mediante TensorFlow.

- **Validación y Generalización:**

La validación de los modelos con conjuntos de datos independientes asegura su generalización y previene el sobreajuste.