

Resumen ejecutivo
Paulina Leal Mosqueda, A01659576
Santiago Nava Figueroa, A01174557

En este documento se presenta un resumen ejecutivo de los resultados obtenidos luego de una indagación de la información otorgada por el equipo de LendSmart. Actualmente manejan un criterio de predicción del 28% en promedio para perfiles default. Como consultores, nuestro objetivo es construir un modelo de clasificación que pueda predecir el estado de su estatus de deuda (Loan Status), para que el equipo de LendSmart pueda tomar decisiones informadas basado en la probabilidad de que un cliente sea de alto riesgo o uno seguro para darle un crédito. Para ello se aplicarán dos modelos de clasificación de análisis discriminante, uno lineal y otro cuadrático, con el objetivo de observar cual es el que brinda mejores resultados para la empresa.

Hallazgos

Una vez que se aplicó el análisis discriminante lineal, se encontraron las 5 variables que más influyen a tener un cliente con deudas default. En este caso, se obtuvieron variables que influyen de manera positiva y de manera negativa, lo que quiere decir que hay elementos que influyen a que la persona tenga una probabilidad menor de tener deudas y viceversa. En particular, cuando una persona tiene un buen historial crediticio, estabilidad laboral y un buen score crediticio, esta suele ser clasificada como alguien que tiene un estatus de préstamo favorable. Por otro lado, si una persona gasta más de lo que tiene en su crédito y si tiene un nivel alto de deuda, entonces, esta se suele clasificar como alguien de alto riesgo. Por ende, un perfil de riesgo, es aquel que mal uso crediticio y una mala relación deuda-ingresos.

Selección de modelo y rendimiento

En base a nuestros modelos construidos y nuestro análisis basado en resultados, podemos decir que en este caso, usar ya sea el modelo QDA o el modelo LDA no representa una disyuntiva con factores contrastantes. Ya que ambos tienen puntajes perfectos en detección de: Aprobación de un buen crédito y Denegación de un mal crédito y no tiene ningún fallo al no suceder aprobaciones de malos créditos y denegaciones de malos créditos.

Esto significa que tomando en cuenta la base de datos original, nuestro modelos son capaces de identificar al 100 por ciento todas las instancias en base a factores subyacentes y las variables predictoras, por lo que internamente, el modelo sabe qué factores tomar en cuenta para predecir un caso y por cómo se maneja esta metodología, si se introducen nuevos datos, la probabilidad de que siga teniendo un buen rendimiento es muy alta.

Es vital tomar en cuenta que si los modelos no fueran perfectos, la decisión óptima sería tomar el modelo que de mejores puntajes f1 score. Ya que este es el que mejor balancea los objetivos de la empresa, ya sea evitar a toda costa las aprobaciones de malos créditos y denegaciones de malos créditos.

A su vez, si se quiere tomar en consideración. La gran ventaja del modelo LDA es el poder visualizar los factores más determinantes a la hora de ejecución del modelo, los cuales se mencionaron anteriormente.

Lo anterior ilustra de manera clara y directa qué factores son los que más pesan de forma objetiva a la hora de clasificar. Destacando la capacidad histórica de pagar, la estabilidad laboral y score crediticio como variables que apoyan a una buena imagen para dar un crédito si son altos y utilización de crédito y el ratio de deuda - ingreso como factores que perjudican la imagen para un ofrecimiento de créditos si son altos.

Por el contrario, el modelo QDA no es fácil de comprender cómo funciona (qué factores toma en cuenta) por lo que no se puede visualizar el cómo se realiza la clasificación, pero esta presenta la ventaja de ser un modelo más eficiente que el LDA. Aproximadamente un 14 por ciento más rápido que LDA en comparación.

Recomendaciones

A partir de estos resultados, se recomienda que la empresa utilice el modelo QDA si lo que se busca es rapidez, pero si el objetivo es la interpretación, se recomienda LDA, ya que, ambos casos presentan los mismos resultados. Seleccionando uno de estos modelos, lo que se busca es asegurar que LendSmart no apruebe préstamos de personas clasificadas como default y que se asegure que evite rechazar préstamos de individuos con un estado sin deuda. Al utilizar este modelo, se estima que la empresa logre evitar el 100% de las personas con estatus default, por lo que, LendSmart lograra reducir su cartera de préstamos actual de la empresa .

Por otra parte, se recomienda que en un futuro se explore más acerca de las características que estos perfiles de alto riesgo tienen, fuera de las variables que esta base de datos proporciona. Esto puede incluir descripciones demográficas, sociales y económicas. Asimismo, es importante mencionar que este tipo de investigación y clasificación debe de evitar tener sesgos discriminatorios que excluyan a un cierto grupo de personas. Por lo tanto, a futuro, a pesar de utilizar el modelo QDA, se debe de verificar que los préstamos que da LendSmart son diversos y no excluye a ningún tipo de persona.