**Propuesta de Proyecto de seminario**

**Presentación del Proyecto**

Sapiencia es una entidad adscrita al distrito de Medellín cuya misión es otorgar créditos condonables o becas a los nacidos o residentes por al menos 3 años en el municipio de Medellín (Medellín, 2023), esto con el fin de garantizar el acceso y permanencia a la educación superior. Uno de los principales dilemas que ha tenido la entidad, es el predecir el costo que puede suponer un beneficiario durante sus estudios postsecundarios, dado la gran variedad de universidades y programas que existen en el Valle de Aburrá. Por ello, se desarrolló un modelo predictivo con el objetivo de mejorar la gestión del recurso al estimar la proyección financiera para cada estudiante en su ciclo académico.

**Objetivo del Proyecto**

Mejorar la gestión del recurso económico al estimar la proyección financiera para cada estudiante en su ciclo académico e identificar qué beneficiarios pueden ser susceptibles de abandonar sus estudios postsecundarios y por ende su crédito o beca en Sapiencia.

**Contexto del problema**

La educación superior abarca todas las modalidades de enseñanza (profesional, técnica, artística, pedagógica, etc.) impartida por universidades, institutos tecnológicos, escuelas normales, etc. (Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO] , 2020, p.14). Estas instituciones están dirigidas a personas que han completado la educación secundaria y tienen como propósito la obtención de un título, grado, certificado o diploma en el ámbito de la educación superior. En Colombia, se contempla la educación superior en los niveles de técnico profesional, y tecnólogo como modalidades de pregrado. A su vez, la educación posgrado comprende los niveles de Especializaciones (relativas a programas de Especialización Técnica Profesional, Especialización Tecnológica y Especializaciones Profesionales), maestrías, doctorados (Ministerio de Educación Nacional, s.f.).

El acceso a la educación superior ha aumentado en todas las regiones del mundo durante las últimas dos décadas (UNESCO , 2020, p.24). Sin embargo, la deserción postsecundaria es uno de los principales desafíos a los que se enfrentan los estudiantes que aspiran a culminar sus estudios universitarios. Ya sea de manera voluntaria, por motivos familiares, económicos, académicos o personales. Esta es una problemática que se presenta a nivel nacional tanto en las Instituciones de Educación Superior (IES) públicas como privadas.

Según lo especifica el SPADIES (sistema para la prevención y análisis de la deserción en la educación superior) un estudiante se clasifica como desertor del sistema de educación superior cuando no se ha matriculado por dos o más periodos consecutivos en algún programa académico. Teniendo esto en cuenta y basándose en datos del SNIES (Sistema Nacional de Información de la educación superior) y el LEE (Laboratorio de Economía y Educación de la Universidad Javeriana) se estima que, entre el año 2000 y 2021, la tasa de deserción se ubicó en un 11% por periodo académico, lo que significa que en promedio cada semestre 1 de cada 10 estudiantes a nivel nacional desertan de sus estudios postsecundarios ya sea de manera temporal o permanente.

Otro obstáculo para la población que quiere acceder a las IES es que no está siendo beneficiada equitativamente, dado que aún existen significativas diferencias en el acceso, especialmente en lo concerniente al nivel de ingresos(UNESCO , 2020, p.9). Un claro ejemplo es que a pesar de identificar un incremento en el acceso a la educación superior durante los últimos 20 años, de los grupos de escasos recursos económicos, siguen estando subrepresentados en estas instituciones (UNESCO , 2020, p.9), es decir, las personas provenientes de entornos con limitaciones financieras tienen una presencia menor en instituciones como universidades o centros educativos superiores en comparación con otros grupos con mayores recursos económicos. Esto sugiere que el principal problema en el financiamiento para personas que realmente necesitan acceder a la educación superior es la falta de una metodología o modelo claro que abarque todos los aspectos socioeconómicos del aspirante.

En Medellín, también se han evidenciado las problemáticas antes mencionadas, es por ello que desde el año 1996 existe, adscrita al distrito de Medellín, la entidad Sapiencia que es la Agencia de Educación Postsecundaria de la Alcaldía de Medellín, encargada de liderar los proyectos y programas de la educación postsecundaria del Distrito de Medellín (Medellín, 2023).

La Agencia de Educación Superior de Medellín - Sapiencia, brinda oportunidades para el acceso y permanencia en la educación superior a estudiantes de bajos recursos económicos, mediante créditos que se pueden condonar total o parcialmente si se cumplen con requisitos basados en méritos académicos y prestación de servicio social (Medellín, 2023). El proceso de inscripción se realiza de manera virtual; con una periodicidad semestral, en la que a cada aspirante se le asigna un puntaje según los criterios de selección establecidos por el reglamento del Fondo. Con el presupuesto disponible, se determina un punto de corte en cada convocatoria, a partir del cual se seleccionan los beneficiarios.

Sin embargo, a lo largo de la trayectoria que lleva la Agencia, se ha identificado una problemática referente a la gestión y ejecución de los fondos destinados a financiar el acceso y permanencia de estudiantes en la educación postsecundaria en Instituciones de Educación Superior públicas y privadas. A pesar de los esfuerzos para proporcionar créditos condonables que faciliten la inclusión educativa, los cálculos financieros y la planificación actual, en ocasiones la cobertura no ha alcanzado a beneficiar a muchos estudiantes de las comunas y corregimientos de Medellín que buscan adelantar o continuar sus estudios postsecundarios. Lo cuál ha generado malestar en los representantes de dichas comunas y corregimientos, dado que, en ocasiones dichos recursos no se llegan a ejecutar ya sea porque el aspirante no cumple con los requisitos necesarios para optar por el crédito o porque a la hora de surtir el proceso de legalización del crédito el recurso ya ha sido consumido por otros aspirantes del fondo que logran legalizar su crédito primero.

Actualmente, la principal problemática radica en la imprecisión en las estimaciones de las proyecciones financieras utilizadas para destinar el recurso en cada convocatoria para las distintas becas y créditos condonables bajo los dos fondos de financiación con los que se cuentan actualmente como lo son el fondos Presupuesto Participativo (PP) y Recurso Ordinario (RO), al igual que la problemática que subyace al estimar la vigencia en la que estos recursos se van a ejecutar que usualmente suele ser proyectada a 12 años.

Adicionalmente, en la entidad se ha observado cierta dificultad para identificar a aquellos beneficiarios que en algún momento a lo largo de su ciclo académico pueden abandonar de manera parcial o total sus estudios. Dado que para la Agencia resulta de vital importancia velar porque los beneficiarios culminen su carrera universitaria y con ello, poder coordinar satisfactoriamente el crédito, es menester hacer un acompañamiento a esos estudiantes con el objetivo de procurar que permanezcan y puedan culminar sus estudios satisfactoriamente.

Para abordar estos desafíos y mejorar la eficacia en la gestión de los fondos, es crucial desarrollar una metodología avanzada y confiable. Un modelo de machine learning, basado en datos históricos de beneficiarios, podría proporcionar estimaciones más precisas sobre la vigencia de los créditos condonables y becas, al igual que, el costo promedio y un modelo auxiliar que permita estimar una probabilidad de deserción en los beneficiarios de manera particular bajo ciertas características de interés. Este modelo permitiría una planificación financiera más exacta y una asignación de recursos más eficiente, optimizando la ejecución de los fondos y garantizando que el apoyo se brinda de manera justa y equitativa a los estudiantes que hayan nacido o residan en Medellín.

**Propuesta de Solución y Alcance**

En Sapiencia, se busca obtener solución al dilema de estimar la proyección financiera de un beneficiario y el costo promedio que puede tener basados en ciertas características muy específicas como el programa, la universidad, la comuna o el estrato. Dado que usualmente este valor se suele estimar de manera empírica basándose en la experiencia adquirida por los respectivos líderes de proyecto en cada convocatoria.

Por lo anterior, la realización de un modelo predictivo usando Machine Learning permitirá tener una medida más precisa de la proyección financiera, a través de la utilización de datos históricos de beneficiarios que han pasado o aún se encuentran dentro de la agencia de educación superior. Donde esencialmente se quiere saber con certeza cuánto puede costar un estudiante en específico que aspira a cierto programa y universidad adscrita al Valle de Aburrá.

Se pretende desarrollar un modelo de machine learning que, al ingresar las características personales de un estudiante, proporcionará estimaciones sobre la duración del crédito y su costo promedio. Este modelo, basado en datos históricos de beneficiarios, permitirá una planificación financiera más precisa y una asignación de recursos más eficiente al anticipar los costos y el tiempo asociado a cada crédito. Al optimizar la gestión de los fondos, no solo mejorará la eficacia en el apoyo a los estudiantes de Medellín, sino que también se garantizará una mayor transparencia y sostenibilidad en los proyectos educativos de Sapiencia. Esto se traduce en impacto positivo en la comunidad educativa.

Adicionalmente, se plantean otros modelos auxiliares que busquen predecir, según otras variables descriptivas del individuo, que tan probable es que este pueda desertar del fondo de beca o crédito condonable, con el objetivo de hacer un seguimiento a fin de evitar los escenarios de deserción estudiantil. Y con ello, cumplir con el deber misional de la entidad, el cuál es garantizar el acceso y la permanencia en la educación postsecundaria.

**Definición del Alcance**

El objetivo principal del proyecto es crear un modelo estadístico que pueda predecir con exactitud el costo total del beneficiario, así como su proyección financiera a lo largo del tiempo. Además, el modelo proporcionará estimaciones sobre el costo promedio que pueden tener los beneficiarios de cierto programa o universidad. Adicionalmente, se tendrán modelos auxiliares que permitan medir la probabilidad que puede existir de que ese estudiante abandone sus estudios y por ende el crédito o beca en Sapiencia. Para la entidad resulta importante tener control sobre estos escenarios ya que esto le permite tomar acción y actuar en concordancia a sus objetivos y misión como entidad prestadora del servicio a los ciudadanos de Medellín.

Se plantea tomar la información de los beneficiarios de los distintos fondos y modalidades tanto de pregrado como de posgrado con corte desde la convocatoria 2019-1 hasta la convocatoria 2024-1, esto dado que es un periodo considerable donde Sapiencia se ha encargado de todo el proceso que implica otorgar becas y créditos condonables.

**Riesgos e Impacto del Negocio**

**Riesgos**

Se identifican como posibles riesgos en la elaboración, desarrollo y consecución del proyecto los siguientes:

* Una estimación errada o sesgada en cuanto a los montos o recursos destinados en las convocatorias a partir del modelo implementado, puede generar proyecciones alejadas a lo que pueda resultar de un proceso de convocatoria. Este escenario, puede prevenirse a través de otros análisis contextuales que puedan dar cuenta de la presencia de variables exógenos como situaciones económicas, políticas, entre otras.
* El sobre ajuste o parametrización del modelo puede ser contraproducente, ya que puede ser susceptible a outliers (valores atípicos) que sobreestiman o subestiman el cálculo de la proyección y la estimación de la proyección. La cual puede ser mitigada a través de la utilización de datos de entrenamiento con el fin de que el modelo puede recalcular los pesos para cada una de las variables regresoras.

**Impactos**

Se identifican las siguientes situaciones de impacto a raíz del desarrollo e implementación del proyecto:

* Puede proporcionar herramientas más sólidas a la entidad a la hora de implementar planes de desarrollo, atracción del público de interés al generar estadísticas y métricas claras respecto a la gestión de proyecciones financieras realizadas sobre los beneficiarios. Como se plantea en estudio acerca de los determinantes de incumplimiento de créditos educativos en Colombia, utilizando el análisis de Statistical Learning y Machine Learning. El objetivo es recopilar la mayor información posible sobre los individuos que solicitan créditos educativos y así poder predecir su futuro comportamiento de pago y default (Granda Rodríguez, 2020).
* Transmitir transparencia y seguridad a la hora de presentar los resultados logrados por Sapiencia al garantizar el acceso y la permanencia a la educación postsecundaria de los ciudadanos en Medellín.
* Conocer la realidad presupuestal de la Agencia, lo cual mejorará las contemplaciones y análisis a la hora de generar CDPs (certificados de disponibilidad de recurso) y la vigencia establecida en los contratos que en definitiva permiten financiar las carreras de los beneficiarios en Sapiencia. El análisis sobre el diseño de los contratos de créditos educativos ayuda en la creación de modelos de referencia como herramienta para identificar y rediseñar las funciones de beneficios y costos de ambas partes (Del Rey y Verheyden, 2009).
* Agilizar y automatizar procesos realizados actualmente en Sapiencia, mediante la implementación de dashboards e informes financieros más concluyentes que puedan orientar a los líderes de proyecto y directores a tomar decisiones más confiables sin basarse tanto en medidas empíricas, que si bien son fruto de la experiencia, en ocasiones dichas estimaciones han generado proyecciones muy alejadas de la realidad de las convocatorias a los distintas becas y créditos condonables.
* Demostrar cómo, la información histórica que se posee de los beneficiarios de los distintos fondos, nos puede servir para predecir los costos y topes presupuestales que se esperan asumir para los posibles beneficiarios de nuevas convocatorias.

**Referencias**

* Medellín, S. (2023a, enero 13). Historia sapiencia - sapiencia. Sapiencia Agencia de Educación Postsecundaria de Medellín. https://sapiencia.gov.co/historia/
* Medellín, S. (2023, 23 noviembre). Mision2 - Sapiencia. Sapiencia Agencia de Educación Postsecundaria de Medellín. <https://sapiencia.gov.co/mision-y-vision-2/>
* Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura [UNESCO]. (2020). Re| Hacia el acceso universal a la educación superior: tendencias internacionales. https://www.iesalc.unesco.org/wp-content/uploads/2020/11/acceso-universal-a-la-ES-ESPANOL.pdf
* §Ministerio de Educación Nacional de Colombia. (2023). Estadísticas de deserción. Sistema de Prevención y Análisis de la Deserción en las Instituciones de Educación Superior (SPADIES).<https://www.mineducacion.gov.co/sistemasinfo/spadies/Informacion-Institucional/357549:Estadisticas-de-Desercion>
* §SPADIES 3.0 - Sistemas información. (s. f.).<https://www.mineducacion.gov.co/sistemasdeinformacion/1735/w3-article-363411.html?_noredirect=1>
* Ministerio de Educación Nacional. (s.f.). *La educación de posgrado*.<https://www.mineducacion.gov.co/>
* Granda Rodríguez, M. A. (2020). *Determinantes del riesgo de incumplimiento en créditos educativos: Un análisis para Colombia* (Tesis de maestría). Universidad EAFIT.<http://hdl.handle.net/10784/24951>