



Automatización y análisis de la Evaluación de Prácticas Académicas en Ingeniería Industrial de la Universidad de Antioquia

Juan Pablo Pérez Atehortua

Trabajo de grado para optar por el título de Ingeniero Industrial

Asesor

Olga Cecilia Usuga Manco, PhD en Ciencias - Estadística

Universidad de Antioquia
Facultad de Ingeniería
Ingeniería Industrial
Medellin, Antioquia, Colombia
2024

Cita	(Pérez Atehortua, 2024)
Referencia	Pérez Atehortua, J. P. (2024). <i>Automatización y análisis de la Evaluación de Prácticas Académicas en Ingeniería Industrial de la Universidad de Antioquia</i> [Trabajo de grado profesional]. Universidad de Antioquia, Medellín, Colombia.

Estilo APA 7 (2020)



Centro de Documentación Ingeniería (CENDOI)

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Tabla de contenido

Resumen	3
Abstract	4
Introducción	5
Planteamiento del problema	9
1 Objetivos	7
1.1 Objetivo general	7
1.2 Objetivos específicos	7
2 Marco teórico	8
3 Metodología	11
4 Resultados	14
6 Conclusiones	23
Referencias	24

Lista de tablas

Tabla 1 Fragmento del resultado de automatización características personales	19
Tabla 2 Fragmento del resultado de automatización competencias y variables textuales	19
Tabla 3 Fragmento de tabla adicional de variables socioeconómicas y académicas	19
Tabla 4 Tabla adicional énfasis académico.	20

Lista de figuras

Figura 1 Diagrama composición PDF	17
Figura 2 Distribución de empresas por macro sector	20
Figura 3 Distribución por énfasis profesional	22
Figura 4 Relación entre variables socioeconómicas y académicas con calificación de competencias	22
Figura 5 Nube de palabras del título de la práctica	24
Figura 6 Nube de palabras de aportes que el estudiante realizo en la empresa	24

Resumen

En el contexto de la Ingeniería Industrial en la Universidad de Antioquia, esta propuesta se centra en la automatización y análisis del proceso evaluativo de las prácticas académicas. Inicialmente, se identifica el proceso de evaluación, que se encuentra en formato PDF por cada estudiante. Luego, se desarrolla un sistema de automatización destinada a un grupo de egresados a lo largo de un año, para extraer la información de estos PDF y convertirla en observaciones de un Dataset. Posteriormente, se realiza un análisis descriptivo para identificar relaciones entre el desempeño en prácticas académicas y factores socioeconómicos y académicos. Además, se analizan las variables textuales, como las mejoras sugeridas por los asesores a los estudiantes, utilizando técnicas de procesamiento de lenguaje natural. Finalmente, se proponen mejoras para el centro de prácticas, los estudiantes, el currículo y el Departamento de Ingeniería Industrial, fundamentadas en los resultados obtenidos del análisis de datos.

Palabras clave: análisis de datos, automatización, practicas académicas, procesamiento de lenguaje natural.

Abstract

In the context of Industrial Engineering at the University of Antioquia, this proposal focuses on the automation and analysis of the evaluative process of academic practices. Initially, the evaluation process, which is in PDF format for each student, is identified. Then, a system of automation is developed targeting a group of graduates over the course of a year to extract the information from these PDFs and convert it into observations of a Dataset. Subsequently, a descriptive analysis is conducted to identify relationships between performance in academic practices and socioeconomic and academic factors. Additionally, textual variables, such as improvements suggested by advisors to students, are analyzed using natural language processing techniques. Finally, improvements are proposed for the practice center, students, curriculum, and the Department of Industrial Engineering, based on the results obtained from the data analysis.

Keywords: data analysis, automation, academic practices, natural language processing.

Introducción

En el ámbito de la Universidad de Antioquia, específicamente en el Departamento de Ingeniería Industrial, se debe abordar de manera integral la evaluación de las prácticas académicas. Este proceso, crucial para el desarrollo de los estudiantes y el fortalecimiento del programa académico, enfrenta diversos desafíos que exigen soluciones innovadoras y efectivas.

El objetivo de este estudio es automatizar y analizar esta evaluación para optimizar su eficiencia y calidad. A partir de la identificación del proceso de evaluación actual, el foco se dirige hacia el desarrollo de un sistema de automatización que facilite la extracción y análisis de la información contenida en los archivos de evaluación, en este caso el objetivo de estudio es un grupo de egresados a lo largo de un año. Además, se busca un análisis descriptivo detallado para comprender las relaciones entre el desempeño en las prácticas académicas y una variedad de factores socioeconómicos y académicos. Este análisis proporcionará perspectivas valiosas para la mejora continua del proceso de evaluación.

Asimismo, se abordará el análisis de las variables textuales presentes en las evaluaciones, incluyendo las sugerencias de mejora brindadas por los asesores a los estudiantes. Este análisis se realizará con técnicas avanzadas de procesamiento del lenguaje natural, para extraer información significativa para tomar decisiones.

Finalmente, se propondrán mejoras concretas para el centro de prácticas, los estudiantes, el plan de estudios y el Departamento de Ingeniería Industrial, a partir de los hallazgos obtenidos del análisis de datos. Estas mejoras estarán respaldadas por evidencia sólida y brindarán una base sólida para la optimización continua del proceso de evaluación en el contexto específico de la Ingeniería Industrial de la Universidad de Antioquia.

Lo que resta del documento se organiza así: primero establece el planteamiento del problema, segundo se definen los objetivos del trabajo, seguido de una revisión de literatura, donde se exploran conceptos, metodologías y teorías consideradas, y el estado del arte que muestra la forma de solucionar la problemática en contextos similares. Luego se expone la metodología empleada para finalmente mostrar y analizar los resultados obtenidos y concluir sobre el trabajo.

Planteamiento del problema

El problema identificado en el Departamento de Ingeniería Industrial de la Universidad de Antioquia se centra en la falta de automatización de los formatos utilizados en el proceso de evaluación de prácticas académicas. Esta carencia impide una retroalimentación oportuna y eficiente durante el proceso de evaluación, lo que limita la capacidad del Departamento para identificar y abordar áreas de mejora de manera efectiva en tiempo real. Aunque los practicantes suelen recibir evaluaciones positivas, es crucial detectar y corregir posibles deficiencias en la evaluación de habilidades que podrían pasar desapercibidas por la falta de una sistemática y automatizada revisión de los datos. Además, se ha observado una ausencia de una caracterización precisa del desempeño de los estudiantes en relación con otras variables socioeconómicas y académicas, lo que dificulta la identificación de patrones y la toma de decisiones informadas para mejorar el proceso educativo. Por tanto, es crucial desarrollar un sistema automatizado de evaluación que permita una retroalimentación eficaz y en tiempo real, así como una caracterización detallada del desempeño estudiantil en relación con otras variables, con el fin de fortalecer la calidad y efectividad de las prácticas académicas en el departamento y contribuir al desarrollo integral de los estudiantes.

1 Objetivos

1.1 Objetivo general

Automatizar y analizar el proceso de evaluación de ejecución de prácticas académicas de Ingeniería Industrial en la Universidad de Antioquia.

1.2 Objetivos específicos

1. Identificar el proceso de evaluación de prácticas académicas en Ingeniería Industrial en la Universidad de Antioquia.
2. Desarrollar un sistema de automatización de los formatos de evaluación de prácticas académicas.
3. Analizar de forma descriptiva las relaciones entre el desempeño en prácticas académicas y diversos factores socioeconómicos y académicos.
4. Analizar las variables textuales presentes en las evaluaciones para identificar mejoras en el desempeño de los practicantes.
5. Proponer mejoras para el centro de prácticas, el currículo y el Departamento de Ingeniería Industrial, basadas en los resultados obtenidos en el análisis previo.

2 Marco teórico

El objetivo de este marco teórico es proporcionar un contexto sólido para comprender los desafíos específicos que enfrenta el proceso de evaluación en el contexto institucional. A través de la revisión crítica de la literatura existente, se busca identificar las mejores prácticas y enfoques que puedan informar la investigación y guiar las propuestas de mejora.

Según Agudelo et al., (2020.), la automatización en el ámbito industrial se refiere al proceso mediante el cual las máquinas realizan tareas sin intervención humana, lo que conlleva a una notable eficiencia en términos de tiempo y costos. Sus orígenes se remontan a épocas prehistóricas, donde se desarrollaron dispositivos simples para reducir el esfuerzo humano, evolucionando gradualmente hacia el uso de fuentes de energía renovable. La introducción de mecanismos de relojería marcó un hito en este progreso, allanando el camino para la implementación de sistemas más complejos a principios del siglo XX, los cuales empleaban elementos mecánicos y electromecánicos para su funcionamiento.

Por otro lado, Hurtado Minotta (n.d.) establece que el análisis estadístico es fundamental en todos los ámbitos de la investigación científica seria, ya que sus técnicas son aplicables en diversas esferas de la vida. Por ejemplo, se emplean encuestas para recopilar datos previos a las elecciones y así anticipar sus resultados. Además, se selecciona aleatoriamente a consumidores para obtener información que permita prever sus preferencias respecto a ciertos productos o servicios. Para quienes toman decisiones en política económica, como asesores presidenciales, ministros y otros altos funcionarios públicos, la estadística es una herramienta invaluable. Los economistas examinan varios indicadores de la situación económica en períodos específicos y emplean esta información para prever su evolución futura. Solamente a través del análisis estadístico es posible adoptar decisiones inteligentes acerca de tasas impositivas, programas sociales, gastos en defensa, políticas laborales e inversiones prioritarias

La estadística, una disciplina que ha evolucionado considerablemente, ha sido tradicionalmente dividida en dos ramas principales: la estadística descriptiva y la estadística inferencial. La primera se enfoca en describir y resumir los datos de manera comprensible, sin buscar generalizaciones más allá del conjunto de datos presente. Por otro lado, la estadística inferencial busca extraer conclusiones sobre poblaciones extensas a partir de muestras disponibles, empleando el cálculo de probabilidades como herramienta fundamental. Históricamente, la estadística se centraba

principalmente en su aspecto descriptivo hasta alrededor del año 1900. No obstante, con los avances en el cálculo de probabilidades y el trabajo pionero de Fisher, Pearson y otros investigadores, la estadística inferencial comenzó a ganar relevancia. Este cambio dio lugar a la aparición de la estadística teórica, que se apartó en cierta medida de los métodos primitivos centrados en la recolección y análisis de datos. En décadas posteriores, especialmente en los años 60, hubo una marcada inclinación hacia la formalización matemática de los modelos inferenciales, a expensas, en ocasiones, de la atención dedicada a los aspectos prácticos del análisis de datos (Batanero & Godino, 2001.).

Adicional, Hidalgo Troya (2019), establece que la mayoría del software estadístico está equipado con herramientas que ofrecen técnicas gráficas listas para usar, mejoradas con medidas estadísticas más elaboradas para su descripción. Estas técnicas permiten explorar las características de la distribución de las variables en análisis, así como las relaciones bivariantes (y multivariantes) entre ellas y el análisis de las diferencias entre grupos. Es crucial recordar que las representaciones gráficas nunca deben suplantar a las medidas de diagnóstico estadístico formales, como los contrastes de ajuste de los datos a una distribución o los contrastes de asimetría. Sin embargo, proporcionan una perspectiva alternativa para comprender la naturaleza de los datos y sus interrelaciones, incluso en casos multivariantes.

Se puede considerar que combinar métodos gráficos con análisis estadísticos detallados puede ofrecer una comprensión más completa y rica de los datos. Las representaciones visuales pueden revelar patrones y tendencias de una manera intuitiva, mientras que las medidas estadísticas proporcionan una validación rigurosa de nuestras interpretaciones. Por lo tanto, integrar ambas aproximaciones puede enriquecer significativamente el proceso de análisis y toma de decisiones.

Al igual que la combinación de métodos gráficos y análisis estadísticos enriquece nuestra comprensión de los datos, la integración de herramientas computacionales, especialmente en el ámbito de la Inteligencia Artificial (IA), desempeña un papel crucial en este proceso. Una de las principales responsabilidades de la IA es el manejo de los lenguajes naturales mediante herramientas computacionales. En este contexto, los lenguajes de programación desempeñan un papel esencial al servir como el vínculo necesario entre los lenguajes naturales y su manipulación por parte de las máquinas. El Procesamiento del Lenguaje Natural (PLN) se define como el empleo de un lenguaje natural para interactuar con las computadoras, requiriendo que estas comprendan las oraciones proporcionadas. La utilización de estos lenguajes naturales facilita tanto el desarrollo

de programas que ejecuten tareas relacionadas con el lenguaje como la creación de modelos para comprender los mecanismos lingüísticos humanos. A pesar de sus ventajas, el uso del lenguaje natural en la comunicación entre humanos y máquinas presenta desafíos únicos que requieren soluciones innovadoras para una interacción fluida y eficaz (Augusto Cortez Vásquez et al., 2009).

La integración de herramientas de procesamiento de lenguaje natural en la inteligencia artificial tiene el potencial de revolucionar la forma en que se interactúa con las máquinas, pero también plantea desafíos significativos en términos de comprensión y comunicación efectiva. Es crucial abordar estos desafíos de manera creativa y colaborativa para aprovechar al máximo el poder transformador de esta tecnología. En este sentido, la utilización de gráficos de nubes de palabras se destaca como una técnica valiosa para analizar datos textuales y comprender patrones emergentes. Estudios recientes, como el realizado por De et al. (2016), son un ejemplo que han demostrado cómo estos gráficos pueden revelar la evolución de temas de investigación a lo largo del tiempo, en un grupo de investigadores por medio de gráficos de nubes de palabras animadas. Se encontró que la presencia destacada de ciertos términos en un año específico, seguida de su desaparición o disminución en años posteriores, puede indicar cambios en los intereses del grupo. Por otro lado, la persistencia de términos similares a lo largo del tiempo sugiere una continuidad en la línea de investigación. Además, se observó que la cantidad de temas o palabras relevantes puede proporcionar información sobre el perfil del grupo, con una constancia en los temas indicando un enfoque conservador y una variabilidad sugiriendo una apertura a nuevos temas.

3 Metodología

La metodología empleada en este estudio es de naturaleza cuantitativa y ha sido estructurada en una serie de etapas esenciales con el propósito primordial de optimizar y analizar el proceso de evaluación de prácticas académicas en el Departamento de Ingeniería Industrial de la Universidad de Antioquia. En esta sección, se exponen los pasos seguidos, organizados en fases, para comprender el proceso evaluativo actual, centrándose en un grupo de egresados durante un año. El objetivo es desarrollar un sistema de automatización eficaz, llevar a cabo análisis estadísticos y descriptivos, analizar variables textuales y proponer mejoras sustanciales basadas en los hallazgos obtenidos. Cada etapa se ha ejecutado de manera metódica y rigurosa, con el fin de obtener resultados confiables y significativos que contribuyan al fortalecimiento del proceso educativo en el departamento. A continuación, se detallan las acciones y procedimientos que han guiado esta investigación, desde la identificación del problema hasta la proposición de soluciones concretas respaldadas por evidencia empírica.

Fase 1: Identificación del proceso de evaluación de prácticas académicas:

En esta primera fase, se llevó a cabo un análisis exhaustivo del proceso de evaluación de prácticas académicas en el Departamento de Ingeniería Industrial de la Universidad de Antioquia, centrándonos en un grupo de egresados a lo largo de un año. El objetivo principal fue identificar y comprender el formato utilizado para la evaluación, específicamente el archivo PDF conocido como "evaluación de desempeño". Este análisis permitió obtener una visión detallada de los criterios y procesos utilizados para evaluar el desempeño de los practicantes en el departamento, sentando las bases para las siguientes etapas de la metodología.

Fase 2: Desarrollo del sistema de automatización:

En esta fase, una vez identificado el proceso de evaluación, se procedió al desarrollo del sistema de automatización utilizando el lenguaje de programación Python y diversas librerías, como Tabula para la extracción de datos de los archivos PDF, y NumPy y Pandas para la manipulación de los datos. El objetivo principal de esta etapa fue automatizar el proceso de extracción de información de los archivos PDF de evaluación, los cuales contienen la información de cada estudiante. Para ello, se implementó un proceso en el cual cada PDF se convierte en una

fila de un Dataset, utilizando las librerías mencionadas. De esta manera, se logró almacenar toda la información de los estudiantes de manera estructurada en un Dataset, lo que facilita su posterior análisis y manipulación. Esta fase fue fundamental para garantizar la eficiencia y precisión del proceso de evaluación de prácticas académicas, sentando las bases para las siguientes etapas de la metodología.

Fase 3: Análisis descriptivo y relaciones.

En esta etapa, se realizó un análisis detallado de los datos obtenidos, además de añadir dos conjuntos de datos adicionales a la información extraída. El primer conjunto de datos contiene información sobre variables socioeconómicas y académicas de los estudiantes, mientras que el segundo conjunto de datos indica el énfasis profesional de cada estudiante, pudiendo ser logística, analítica o finanzas, sin ser mutuamente excluyentes. Se procedió a realizar una limpieza exhaustiva de la información para garantizar la calidad de los datos.

Posteriormente, se llevaron a cabo diferentes combinaciones de gráficos utilizando la librería Matplotlib, con el fin de identificar y visualizar las relaciones mencionadas anteriormente entre el desempeño en prácticas académicas y los diversos factores socioeconómicos, académicos y profesionales de los estudiantes. Este análisis permitió obtener resultados significativos y reveladores que contribuyeron a comprender mejor el proceso evaluativo y sus resultados, sentando las bases para las siguientes fases de la metodología.

Fase 4: Análisis de variables textuales.

En esta fase, se emplearon diversas librerías de procesamiento de lenguaje natural, como NLTK y Spacy para preprocesar y limpiar el texto contenido en las variables textuales de las evaluaciones de desempeño de los estudiantes. El objetivo principal fue preparar los datos textuales para su posterior análisis y visualización.

El enfoque principal de este análisis se centró en la generación de gráficos de nubes de palabras, utilizando la librería Wordcloud. Estos gráficos permitieron identificar las palabras más repetidas en las variables textuales, lo que a su vez proporcionó insights sobre los aspectos más destacados en las evaluaciones. Este análisis de palabras clave fue fundamental para comprender la importancia y relevancia de ciertos aspectos señalados por los asesores en las evaluaciones de desempeño de los estudiantes.

Fase 5: Propuesta de mejoras.

Basándose en los resultados obtenidos del análisis de datos, se propusieron mejoras específicas para el centro de prácticas, los estudiantes, el currículo y el Departamento de Ingeniería Industrial. Estas mejoras fueron diseñadas para abordar las deficiencias identificadas en el proceso evaluativo y mejorar la calidad global de la educación en Ingeniería Industrial en la Universidad de Antioquia.

Cada fase de la metodología fue realizada de manera sistemática y rigurosa, con el objetivo de obtener resultados confiables y significativos que contribuyan al fortalecimiento del proceso educativo en el Departamento de Ingeniería Industrial.

4 Resultados

A continuación, se presentan los resultados obtenidos a lo largo de este estudio, iniciando en la identificación del proceso de evaluación hasta la propuesta de mejoras concretas para el Departamento y los estudiantes.

Identificación del proceso de evaluación de prácticas académicas

Se identificó cómo estaban compuestos los archivos en los cuales se encuentra la información de las prácticas académicas de cada estudiante. Cada estudiante cuenta con 1 carpeta con cuatro archivos, los cuales son, acta de inicio, acta de finalización, evaluación al escenario de prácticas y evaluación de desempeño del estudiante. Explorando estos archivos, se establece que el archivo de referencia para nuestro trabajo investigativo es aquel que presenta el nombre "Evaluación_desempeño_estudiante" ('estudiante' varía dependiendo el nombre del estudiante) y está en formato PDF, el cual contiene información del estudiante, las calificaciones de las competencias y características personales otorgadas por el asesor a los practicantes. Estas calificaciones están clasificadas como E (Excelente), B (Bueno), A (Aceptable) y D (Deficiente), así como diversas variables textuales. Estas variables incluyen, entre otras, las oportunidades de mejora pactadas entre el asesor y el estudiante, comentarios adicionales del asesor sobre el desempeño del practicante y cualquier otra información relevante para la evaluación del desempeño en las prácticas académicas. La composición del PDF se observa en la **Figura 1**. Este análisis permitió obtener una visión detallada de los criterios y procesos utilizados para evaluar el desempeño de los practicantes en el departamento.

Figura 1

Diagrama composición PDF.



Desarrollo del sistema de automatización

Durante el desarrollo de esta fase, se logró implementar un sistema de automatización efectivo para extraer información vital de archivos PDF relacionados con evaluaciones de estudiantes. Este proceso se llevó a cabo mediante el uso de Python y varias librerías especializadas, incluyendo Tabula para la extracción de datos de archivos PDF y NumPy y Pandas para su manipulación y procesamiento.

El sistema de automatización desarrollado fue diseñado para enfrentar el desafío de procesar una gran cantidad de archivos PDF de evaluación de manera eficiente. A través de una cuidadosa manipulación de datos, se logró establecer un flujo de trabajo robusto que permitió la extracción y estructuración de datos de manera precisa y rápida. En primer lugar, se realizaron los pasos necesarios para preparar el entorno de desarrollo, que incluyeron la configuración de la ruta principal y la identificación de las carpetas de evaluaciones dentro de ella. Posteriormente, se iteró sobre estas carpetas, buscando archivos que contuvieran la palabra clave "desempeño", lo que indicaba la presencia de los documentos relevantes. Una vez localizados los archivos pertinentes, se procedió a la extracción de datos utilizando la librería Tabula. Esta herramienta demostró ser crucial en la tarea de convertir el contenido de los archivos PDF en datos estructurados y manipulables. Durante este proceso, se realizaron múltiples operaciones de preprocesamiento para limpiar y organizar los datos, lo que garantizó la coherencia y calidad del conjunto final de información. Entre las operaciones realizadas se incluyen la eliminación de columnas innecesarias, la corrección de valores erróneos y la estructuración adecuada de los datos para su posterior análisis. Además, se llevaron a cabo búsquedas específicas dentro de los documentos PDF para extraer información relevante, como datos personales de los estudiantes y detalles de sus evaluaciones.

El resultado de este proceso fue la creación de un Dataset completo y estructurado (**Tabla 1, Tabla 2**), que contiene toda la información relevante extraída de los archivos PDF. Cada entrada en este conjunto de datos representaba la información completa obtenida de un estudiante en particular, lo que facilitó su análisis posterior.

Tabla 1*Fragmento del resultado de automatización características personales*

Nombre completo del estudiante:	Documento de Identidad:	Nombre de la empresa:	Fecha de inicio:	Fecha de finalización:	Título de la práctica:
Estudiante X		Línea Directa S.A.S	16/06/2022	15/12/2022	Análisis en los procesos de gesti
Estudiante Y		Vicerrectoría de Investigación	23-Mayo-2022	27-Septiembre-2022	Sistematización de Experiencias
Estudiante Z		Universidad de Antioquia	25/07/2022	24/01/2023	Identificación, documentación y

Tabla 2*Fragmento del resultado de automatización competencias y variables textuales*

Seguridad de sí mismo	Puntualidad	Presentación (imagen)	Orden en el trabajo	Actitud Proactiva	Oportunidades de mejoramiento	Aportes que el estudiante realizó
A	E	E	E	B	Análisis en los procesos de gestión	Diseñar y construir herramientas p
E	E	E	E	E	Ninguna	Aportar de manera rigurosa al pro
E	E	E	E	E	Continuar fortaleciendo sus capac	Estandarización y mejoramiento d

Análisis descriptivo y relaciones

Con el objetivo de llevar a cabo un análisis comparativo exhaustivo del desempeño del practicante, se ha ampliado el Dataset original con dos tablas adicionales: **Tabla 3 y Tabla 4**. Estas nuevas tablas permiten explorar la relación entre las variables académicas, socioeconómicas y el énfasis profesional del practicante con el desempeño del estudiante en las practicas académicas. Es importante destacar que los énfasis no son mutuamente excluyentes, lo que significa que un estudiante puede participar en más de un énfasis al mismo tiempo. En esta tabla correspondiente al énfasis, la 'x' representa el énfasis que está realizando el estudiante.

La adición de estas tablas es crucial para comprender mejor cómo factores como el entorno académico y socioeconómico del practicante pueden influir en su desempeño, así como también cómo el énfasis profesional elegido puede tener un impacto en la evaluación final proporcionada por el asesor. Al unir estos conjuntos de datos, se busca identificar posibles relaciones y patrones que puedan contribuir a una comprensión más profunda de los factores que influyen en el éxito del practicante

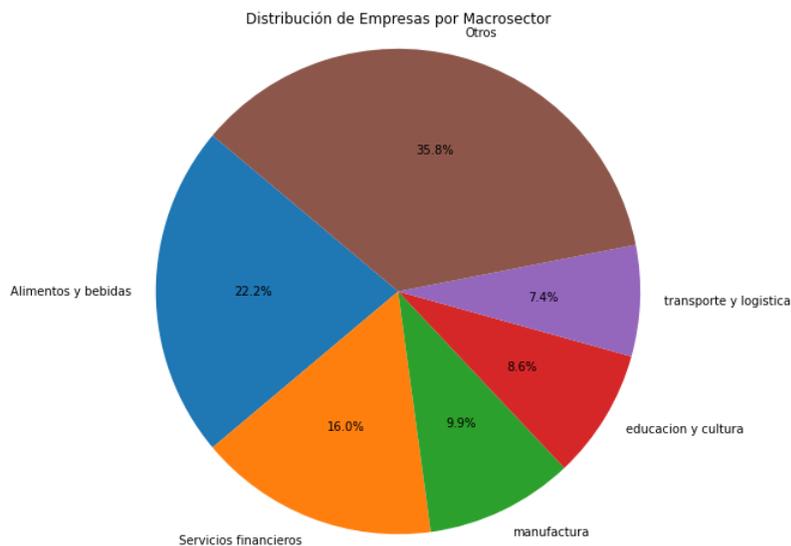
Tabla 3*Fragmento de tabla adicional de variables socioeconómicas y académicas.*

RANGO	PERIODOS_PRUEBA	TIPO_ACEPTACION	NOMBRE_COLE	NATURALEZA	CREDGRADO	CREDAPROB	NUMSEMEST
NORMAL	0	INDIGENA	INSTITUCION EDUCATIVA JOSE CELESTINO MUTIS	OFICIAL	160	167	12
NORMAL	2	POR-EXAM	COLEGIO DOCTOR JOSE MANUEL RESTREPO VELEZ	OFICIAL	145	174	21
SOBRESALIE	1	POR-EXAM	CENTRO FORMATIVO DE ANTIOQUIA CEFA	OFICIAL	160	152	13
SOBRESALIE	0	POR-EXAM	INEM JULIAN MOTTA SALAS	OFICIAL	160	155	15
SOBRESALIE	0	POR-EXAM	COLEGIO SAN VICENTE DE PAUL	OFICIAL	160	163	10
NORMAL	0	POR-EXAM	ESCUELA NORMAL SUPERIOR DE JERICO	OFICIAL	160	173	12

Tabla 4*Tabla adicional énfasis académico.*

Nombre	CC	LOGISTICA	ANALITICA	FINANZAS
Estudiante X		X		
Estudiante Y				X
Estudiante Z		X		

Con el propósito de consolidar la información y simplificar el análisis, con el Dataset original se fusionaron los conjuntos de datos adicionales en uno único, con el cual se trabajó de manera interactiva. Uno de los objetivos fue identificar los sectores en los que los estudiantes de Ingeniería Industrial estaban involucrados laboralmente. Para ello, se elaboró un gráfico circular que revela la distribución de sectores representativos. De la **Figura 2** se observa que, entre los sectores identificados, se destaca un alto porcentaje de empresas relacionadas con alimentos y bebidas, como Nutresa y Postobón. Además, se observa que un segundo porcentaje significativo corresponde a entidades del sector financiero, como Bancolombia y Davivienda.

Figura 2*Distribución de empresas por macro sector.*

A demás, se graficó la frecuencia del énfasis en la **Figura 3** para analizar la distribución de los estudiantes en cada área de especialización. Esta visualización nos permite identificar patrones y tendencias en las preferencias profesionales de los estudiantes, proporcionando un punto de

partida crucial para nuestro análisis más amplio. También, para abordar la evaluación de las competencias de los estudiantes, se realizaron gráficos que exploran todas las combinaciones posibles entre las variables socioeconómicas y académicas, el énfasis profesional y las calificaciones de competencias y características personales del estudiante en las prácticas académicas. Durante este proceso, se identificaron numerosas instancias donde los estudiantes fueron calificados como excelentes, sin embargo, también se observaron casos donde las calificaciones fueron catalogadas como buenas. Esta variabilidad en las calificaciones destacó áreas con oportunidades de mejora **Figura 4**.

Al analizar estos datos, se evidencia claramente que hay ciertos patrones y tendencias que merecen atención. Según la **Figura 3**, la mayoría de los estudiantes optan por enfocarse en el énfasis analítico, seguido de finanzas y logística. Además, es importante destacar que los énfasis no son mutuamente excluyentes, y un 1.30% de la población realizó énfasis en analítica y logística simultáneamente.

Al examinar la **Figura 4**, se observa que, en cuanto a la comunicación oral por énfasis, los estudiantes calificados como B (Bueno) en el énfasis analítico representan una proporción más alta que aquellos calificados como E (Excelente). Esto sugiere oportunidades de mejora en la comunicación oral para el énfasis de analítica. Asimismo, al analizar la comunicación oral por sexo, se aprecia que la proporción de calificaciones B es más alta que las calificaciones E para el género masculino, indicando posibles áreas de mejora en la comunicación oral para los hombres.

También en la **Figura 4** en relación con los gráficos de la capacidad de decisión y la seguridad en sí mismo por estrato, se identifica que en el estrato 3, la proporción de calificaciones B es menor que la proporción de calificaciones E. Aunque la proporción E es mayor, esta discrepancia podría generar una alerta, ya que indica una diferencia reducida entre ambas calificaciones.

A partir de lo anteriormente expuesto, se evidencia que existen más falencias en habilidades blandas que en conocimientos específicos. Por lo tanto, es crucial abordar estas áreas de mejora para fortalecer el perfil del estudiante. En conclusión, es necesario enfocarse en el desarrollo de habilidades blandas complementarias a los conocimientos técnicos para garantizar un desarrollo integral del estudiante.

Figura 3

Distribución por énfasis profesional.

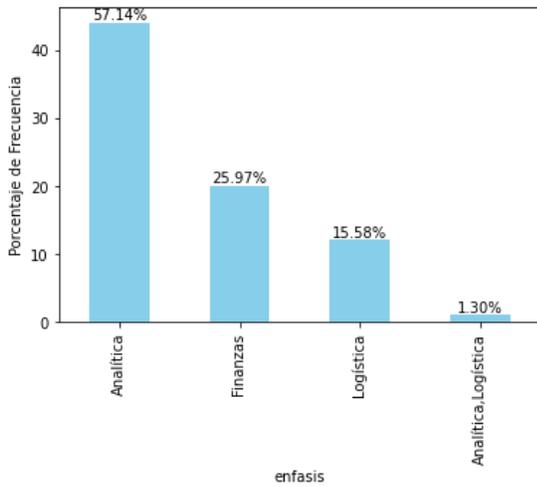
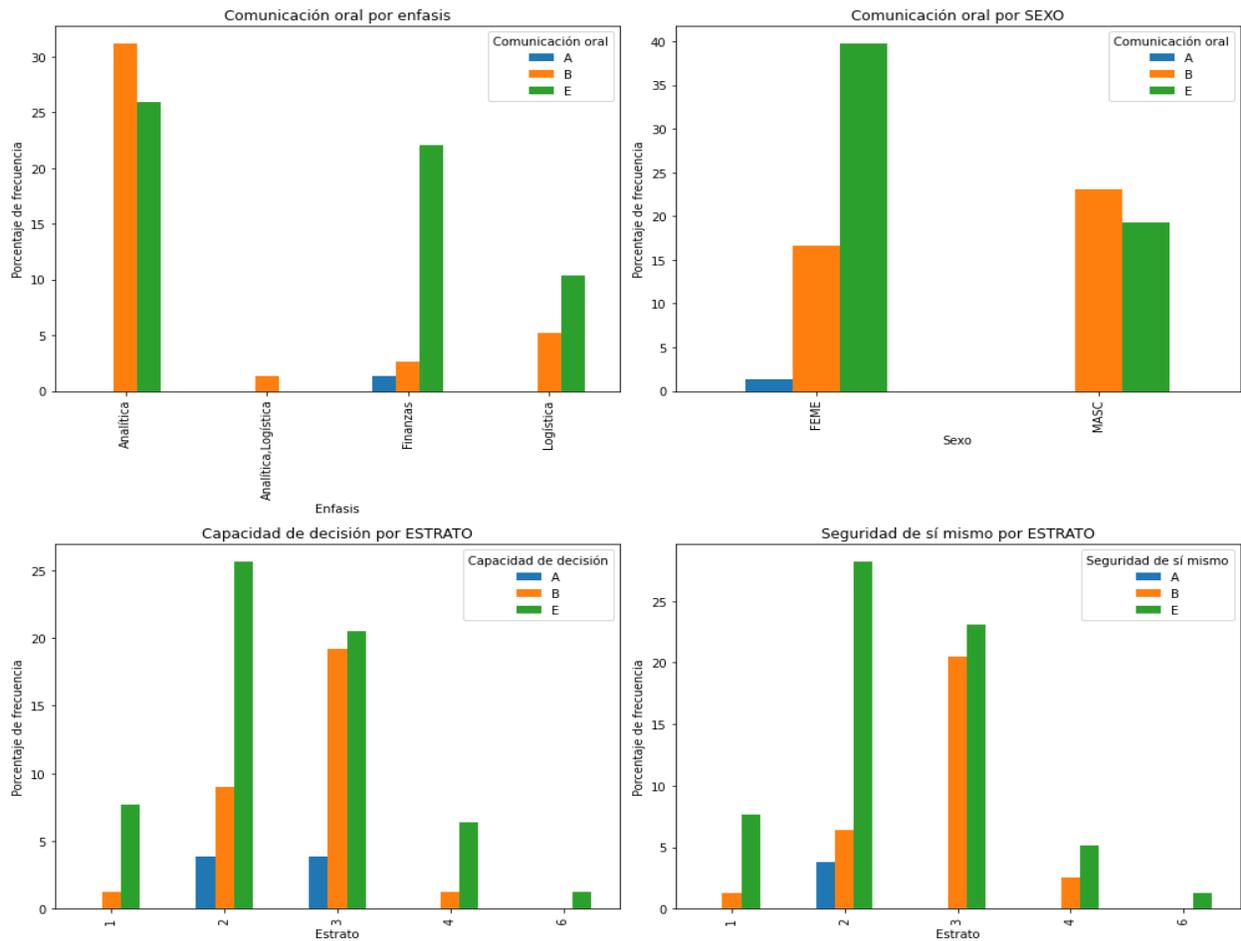


Figura 4

Relación entre variables socioeconómicas y académicas con calificación de competencias.



Análisis de variables textuales

De la automatización se evidenció que había variables textuales de las cuales se podría obtener información relevante y que se podrían usar las nuevas tecnologías como el procesamiento de lenguaje natural para lograr este fin.

El propósito principal es visualizar las palabras más frecuentes en los comentarios textuales de los asesores y en los títulos de las prácticas, lo que nos permite identificar los temas dominantes en las experiencias de práctica de los estudiantes de ingeniería industrial. Esta información es fundamental para comprender mejor las áreas de enfoque, los puntos fuertes y débiles de los estudiantes, así como las tendencias generales en las evaluaciones de prácticas. Al obtener las palabras más frecuentes en los comentarios y los títulos, podemos identificar los temas recurrentes, las habilidades más valoradas y las áreas que podrían necesitar mayor atención o desarrollo.

En esta etapa se identificaron las variables textuales y se logró crear una función llamada `generate_wordcloud` que se utilizó para crear una nube de palabras a partir de Dataset y una columna específica del mismo. El propósito principal es visualizar las palabras más frecuentes en el texto contenido en esa columna, lo cual es fundamental para comprender mejor el contenido y las tendencias dentro del conjunto de datos. Primeramente, se extrae el texto de la columna, se convierte a minúsculas y se elimina cualquier valor nulo. Luego, se limpia el texto eliminando caracteres que no son letras y palabras individuales. Se utilizan las bibliotecas `re`, `pandas`, `nlk` y `Spacy` para estos procesos. Posteriormente, se eliminan las palabras 'basura' (stopwords) en español, que son palabras comunes y poco significativas. Además, se lematizan las palabras restantes, es decir, se las convierte a su forma base, por ejemplo, la palabra `manejo`, `manejando` o `maneje`, quedaría en `manejar`. Finalmente, se genera la nube de palabras utilizando la biblioteca `WordCloud` y se muestra la visualización.

En la **Figura 5** se realizó el gráfico de nube de palabras para evidenciar cual era la problemática más frecuente en el que los estudiantes de Ingeniería Industrial estaban realizando la práctica, se logra evidenciar que la palabra con más peso o más grande es, "proceso", sugiere que este tema es central en las actividades de práctica de los estudiantes. Además, las palabras "gestión" e "implementación" también son destacadas, lo que indica que estas áreas temáticas son importantes y están relacionadas con el proceso en cuestión. Esto sugiere que los estudiantes están involucrados en actividades relacionadas con la gestión e implementación de procesos en el

Propuesta de mejoras

Con base a los resultados evidenciados en este trabajo se presentan un par de mejoras para el departamento de ingeniería industrial

- El documento donde se evidencia la calificación de desempeño del asesor hacia el practicante debe estar siempre en el mismo formato, se recomienda formato PDF para la utilización de la automatización creada para este fin. Adicional a esto que el documento presente el mismo nombre, ya que se evidencio que variaba mucho de un estudiante a otro. Por último, que se exija el adecuado diligenciamiento del formato, ya que se evidenciaban algunos formatos llenados erróneamente y esto dificulta tanto la interpretación como la extracción de la información.
- Con base en lo evidenciado en la comparación de variables socioeconómicas y académicas con la calificación del asesor. Se muestran falencias en comunicación oral, seguridad en sí mismos y planificación y manejo del tiempo en la comunidad universitaria, es recomendable implementar una serie de acciones para mejorar estas áreas. Se sugiere incrementar programas de capacitación y talleres prácticos que fomenten habilidades de comunicación oral, incluyendo técnicas de presentación y expresión verbal. Además, se pueden ofrecer sesiones de desarrollo personal y coaching para fortalecer la confianza y autoestima de los estudiantes, así como herramientas para mejorar la planificación y gestión del tiempo, como técnicas de organización y establecimiento de metas claras. Es esencial proporcionar recursos y apoyo continuo para garantizar la efectividad de estas iniciativas y promover un ambiente universitario que fomente el crecimiento personal y académico.
- Después de analizar la nube de palabras y observar la importancia de los términos "indicadores" y "tableros", se recomienda enfocar los esfuerzos en el desarrollo de cursos especializados en la creación y análisis de indicadores clave de rendimiento (KPIs) y en el diseño y utilización de tableros de control. Estos cursos pueden abordar temas como la selección adecuada de indicadores relevantes para evaluar el desempeño en diferentes áreas, así como técnicas avanzadas para la visualización y presentación de datos en tableros efectivos.

6 Conclusiones

El sistema de automatización desarrollado ha facilitado una gestión más eficiente de las evaluaciones de los estudiantes, logrando reducir significativamente el tiempo necesario para procesarlas. Este hito no solo resalta el poder y la versatilidad de las herramientas de programación, sino que también ofrece una solución práctica y escalable para futuras tareas similares en el ámbito educativo y más allá.

Además, los análisis llevados a cabo en este estudio son de suma importancia debido a su capacidad para identificar áreas críticas de mejora y oportunidades de desarrollo en la formación académica y preparación profesional de los estudiantes de ingeniería industrial. Desde la detección de habilidades blandas necesarias para el éxito laboral hasta la visión integral que proporcionan, estos análisis orientan al departamento hacia la implementación de medidas concretas que fortalezcan el perfil estudiantil y los preparen para enfrentar con eficacia los desafíos del mundo laboral actual y futuro. Asimismo, no solo ofrecen una evaluación detallada de la situación actual, sino que también sirven como una base sólida para la toma de decisiones informadas que promuevan el crecimiento académico y profesional de los estudiantes, asegurando así una formación de calidad y una preparación adecuada para su éxito en la carrera profesional.

Además, para futuros trabajos relacionados con este estudio, se sugiere ampliar el alcance temporal y considerar la inclusión de datos de múltiples jornadas académicas. Dado que la investigación actual se centró exclusivamente en las jornadas académicas del año 2023, incorporar información de otras ediciones podría brindar una perspectiva más completa y robusta de los resultados, permitiendo una evaluación más exhaustiva de las tendencias y patrones observados.

Referencias

- Agudelo, N., Tano, G., & Andrés Vargas, C. (2020). *HISTORIA DE LA AUTOMATIZACIÓN*.
- Augusto Cortez Vásquez, M., Hugo Vega Huerta, M., Jaime, L., & Quispe, P. (2009). *Procesamiento de lenguaje natural*.
- Batanero, C., & Godino, J. D. (2001). *ANÁLISIS DE DATOS Y SU DIDÁCTICA Profesores*.
- De, F., Castillo, L., & Santos, A. S. (2016). *Nubes de palabras animadas para la visualización de información textual de Publicaciones Académicas*. <https://tagul.com/>
- Hidalgo Troya, A. (2019). *REVISTA SIGMA* (Issue 1). <http://coes.udenar.edu.co/revistasigma/articulosXV/1.pdf>
- Hurtado Minotta, E. A. (n.d.). *curso-de-estadistica-descriptiva-y-analitica_opt*.