
	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
---	---	---

Caso a trabajar

La Organización de las Naciones Unidas (ONU) adopta, el 25 de septiembre del año 2015, la Agenda 2030ⁱ para el desarrollo sostenible, cuyo fin es reducir la pobreza, garantizar acceso a la salud y educación, buscar igualdad de género y oportunidades, disminuir el impacto ambiental, entre otros. Esta agenda se basa en 17 objetivos de desarrollo sostenibleⁱⁱ (ODS) y 169 metas (derivadas de los diferentes ODS).



Dentro del trabajo en conjunto de diferentes entes para alcanzar el cumplimiento de los ODS, muchas entidades tienen como enfoque el seguimiento y la evaluación de las políticas públicas y su impacto a nivel social. Este es el caso del Fondo de Poblaciones de las Naciones Unidas (UNFPAⁱⁱⁱ) que, junto con entidades públicas y haciendo uso de diferentes herramientas de participación ciudadana, busca identificar problemas y evaluar soluciones actuales, relacionando la información con los diferentes ODS. En este contexto, uno de los procesos que requiere de un mayor esfuerzo es la clasificación de la información textual que es recopilada, ya que es una tarea que consume gran cantidad de recursos y para la cual se requiere un experto. Es por esto que, en los últimos años, UNFPA ha venido trabajando, en conjunto con la Universidad de los Andes, en la implementación de diferentes estrategias de clasificación de textos, que les permitan hacer un análisis automatizado de opiniones que representan la voz de los habitantes locales sobre problemáticas de su entorno particular.

Para apoyar a UNFPA en este proceso se ha planteado el **primer proyecto del curso**, cuyo **objetivo es desarrollar un modelo de clasificación, con técnicas de aprendizaje automático, que permita relacionar de manera automática un texto según los ODS**. Al igual que **desarrollar una aplicación que facilite la interacción con el resultado de dicho modelo**. El modelo podrá ser utilizado entonces para la interpretación y análisis de la información textual que es recopilada a través de diferentes fuentes por UNFPA en procesos de planeación participativa para el desarrollo a nivel territorial.

En particular se van a abordar dos problemas similares, de los cuales, cada grupo dependiendo de su número asignado en BloqueNeón, tendrá asociado uno:

- Objetivos de desarrollo sostenible 3, 4 y 5.** Grupos con identificador impar.
- Objetivos de desarrollo sostenible 6, 7 y 16.** Grupos con identificador par.

Pasos previos

 <p>Universidad de los Andes Colombia</p> <p>Acreditación institucional de alta calidad 10 años Mineducación Presidencia UIS 9 de enero de 2023</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

- Planear la interacción con un grupo interdisciplinario para identificar usuarios y posibles herramientas a desarrollar para la interacción del resultado del modelo desarrollado.

Descripción de la entrega



A continuación, se describen cada uno de los entregables solicitados en esta etapa.

I. Documento que describa cómo fue aplicada la metodología de analítica de textos y los resultados obtenidos en cada una de las fases:

(10%) Entendimiento del negocio y enfoque analítico. Definición de los objetivos y criterios de éxito desde el punto de vista del negocio. En este punto debe incluir qué es un ODS, en particular describir en detalle, los ODS involucrados en el proyecto asignado y el impacto que puede tener en Colombia el hecho de ser abordados. Descripción del enfoque analítico para alcanzar los objetivos del negocio. Finalmente, incluya el correo de los estudiantes del curso de estadística con los cuales hizo contacto para trabajar la segunda etapa de este proyecto e indique la fecha en que se reunirá con ese contacto y el canal que utilizará para mostrar el resultado de esta etapa e iniciar el trabajo de la etapa2.

Incluya los elementos solicitados en la siguiente tabla:

Oportunidad/problema Negocio	
Enfoque analítico (Descripción del requerimiento desde el punto de vista de aprendizaje automático) e incluya las técnicas y algoritmos que propone utilizar.	

	Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20	
---	---	---

Organización y rol dentro de ella que se beneficia con la oportunidad definida	
Contacto con experto externo al proyecto	



(20%) Entendimiento y preparación de los datos. Perfilamiento y análisis de la calidad de los datos. Además, tratamiento de los datos (preparación o transformaciones requeridas), de acuerdo con el dominio, las técnicas y los algoritmos seleccionados para resolver las tareas.

(25%) Modelado y evaluación. Aplicación de **mínimo tres algoritmos** diferentes para la tarea de aprendizaje automático seleccionada. En esta parte deben describir los algoritmos utilizados para la construcción de los modelos y presentar los resultados de la evaluación cuantitativa. Adicionalmente, deben justificar la selección del modelo, mostrando los diferentes modelos probados e incluir los nombres de los estudiantes que trabajaron en cada uno.

Al igual que en los laboratorios, en la nota habrá un porcentaje específico para el aporte individual de cada miembro del grupo del proyecto.



(15%) Resultados. Descripción de los resultados obtenidos, que permita a la organización comprenderlos, haciendo énfasis en el análisis de las métricas de calidad arrojadas por los modelos utilizados y cómo aportan en la consecución de los objetivos del negocio. Incluir posibles estrategias que la organización debe plantear relacionadas con los resultados obtenidos en los modelos y una justificación de por qué esa información es útil para ellos. Este resultado debe estar en el documento y deben generar un video de **máximo 5 minutos** explicando su proyecto y resultados. El video debe ser publicado en el padlet respectivo (ver condiciones de entrega).

Importante: justifique cada decisión tomada en cada una de las fases.

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	--	---

(10%) Mapa de actores relacionado con un producto de datos creado con el modelo analítico construido. Descripción de una organización que puede beneficiarse del resultado del modelo analítico planteado. Dentro de la organización identifique posibles tipos de actores relacionados con el producto tales como clientes/financiadores/proveedores/beneficiados/etc. y plantee un beneficio y un riesgo que puede tener dicho tipo de actor. Por ejemplo, en el contexto de un modelo de deserción de estudiantes en una universidad, este entregable podría ser el siguiente:



Rol dentro de la empresa	Tipo de actor	Beneficio	Riesgo
Coordinación académica	Usuario-cliente	Apoya la labor de consejería preventiva de los estudiantes	Si el modelo no tiene un buen desempeño, puede estar alertando a estudiantes que no van a desertar o dejando de lado los que sí.
Dirección de ingeniería de sistemas y computación	Financiador	Mecanismo de toma de decisiones basado en información, menos sensible a la intuición del personal	En caso de que el modelo no funcione es dinero mal invertido y pudo dejarse de hacer un proyecto con mayor impacto y viabilidad.
Dirección de servicios de tecnologías de información	Proveedor	Garantiza el cumplimiento de estándares de calidad de los productos desarrollados, que incluye elementos como la seguridad y privacidad de los datos utilizados.	Manejo incorrecto de los datos que lleve a la violación de la privacidad de los datos.
Estudiante	Beneficiado	Recibe una consejería a tiempo que le permite buscar las ayudas que pueden impedir su deserción del programa.	Afectación psicológica que lo lleva a bajar el nivel académico dado que fue alertado o acompañado de forma incorrecta.

 <p>Universidad de los Andes Colombia</p> <p>Accreditación institucional de alta calidad 10 años Ministerio de Educación Presidencia del S.E. 9 de enero de 2020</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

Ten presente esta frase de Albert Einstein: "La preocupación por el hombre y su destino debe ser siempre el principal interés de todos los esfuerzos técnicos. Nunca olvides esto en medio de tus diagramas y ecuaciones".

(8%) Trabajo en equipo

- Describir los roles y las tareas realizadas por cada integrante del grupo (ver sugerencias en la siguiente viñeta), los tiempos, en número de horas, dedicados para realizar lo que le fue asignado, al igual que el algoritmo trabajado, los retos enfrentados en el proyecto y las formas planteadas para resolverlos. Adicionalmente, de acuerdo con la planeación establecida para realizar el proyecto y la realidad en el desarrollo del mismo, reflexionen sobre la forma como repartiría 100 puntos entre los integrantes del grupo e identifique puntos a mejorar para la siguiente entrega del proyecto.
- A nivel del trabajo en equipo se sugieren los siguientes roles para cada integrante. Un integrante puede desempeñar más de un rol.
 - **Líder de proyecto:** Está a cargo de la gestión del proyecto. Define las fechas de reuniones, pre-entregables del grupo y verifica las asignaciones de tareas para que la carga sea equitativa. Se encarga de subir la entrega del grupo. Si no hay consenso sobre algunas decisiones, tiene la última palabra.
 - **Líder de negocio:** Es responsable de velar por resolver el problema o la oportunidad identificada y estar alineado con la estrategia del negocio para el cual se plantea el proyecto. Debe garantizar que el producto se puede comunicar de forma apropiada. Debe encargarse de contactar al grupo de expertos de estadística para determinar la fecha en la que se van a reunir para revisar los resultados de esta etapa e iniciar el trabajo de la etapa 2.
 - **Líder de datos:** Se encarga de gestionar los datos que se van a usar en el proyecto y de las asignaciones de tareas sobre datos. Debe dejarlos disponibles para todo el grupo.
 - **Líder de analítica:** Se encarga de gestionar las tareas de analítica del grupo. Se encarga de verificar que los entregables cumplen con los estándares de análisis y que se tiene el “mejor modelo” según las restricciones existentes.
- Se recomiendan al menos las siguientes reuniones de grupo:
 - **Reunión de lanzamiento y planeación:** Para definir roles y forma de trabajo del grupo. Se genera lluvia de ideas sobre la forma de resolver el proyecto.

	Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20	
---	---	---



- **Reunión de ideación:** Una vez se han explorado los datos del proyecto, la reunión de ideación busca definir la organización/empresa/institución y el rol dentro de ella, que se beneficia de la solución analítica que van a desarrollar.
 - **Reuniones de seguimiento:** Se recomienda mínimo una reunión de seguimiento semanal corta. También pueden ser correos de avance según lo defina el grupo. Pueden tener un tablero de control de las tareas, utilizando herramientas como Trello.
 - **Reunión de finalización:** Para consolidar el trabajo final, verificar el trabajo del grupo y analizar los puntos a mejorar para la siguiente etapa del proyecto.
- (12%) Sustentación y evaluación del aporte individual. En este punto debe entregar un archivo csv, con el mismo formato del conjunto de datos inicial, para el cual adiciona la columna con el ODS, que le asignó el modelo analítico construido. Este archivo se utilizará para determinar el f1-score de su modelo y compararlo con los de los diferentes grupos para asignar parte de la nota del grupo.

II. Conjunto de datos resultado de la fase de entendimiento y preparación de los datos. Requerido para realizar la sustentación.

III. Se debe crear un repositorio para el proyecto del curso, en el cual debe crear una wiki para incluir los diferentes entregables. Debe almacenar el código de la solución creada y referenciarlo en los entregables. Puede acceder a beneficios de estudiante en <https://education.github.com/pack>. El entregable en BloqueNeón es la dirección de la wiki (dentro del repositorio), donde registró todo lo solicitado en esta etapa del proyecto. El github de su proyecto puede ser un repositorio público o compartido a través de githubfront para garantizar que los monitores tengan acceso.

Condiciones de la entrega

- El proyecto se realiza en grupos de mínimo 2 estudiantes y máximo 3. Los estudiantes pueden estar en distintas secciones.

	Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20	
---	---	---

- El documento a entregar tiene máximo 10 páginas (sin incluir portada, tabla de contenido, ni referencias), a una columna y con letra arial, tamaño 12.
- Deben publicar los resultados del proyecto usando el video generado, en el padlet <https://uniandes.padlet.org/mavillam/exposicion-proyecto-analitica-de-texto-de-bi-202320-6s23hgij5y5tb29h> con clave: ProyBI232
- Deben entregar los modelos analíticos desarrollados (Software) y todo lo relacionado con el proyecto para poder ejecutarlos el día de la sustentación.
- Deben entregar el video o en enlace al mismo
- Se sustenta el proyecto y se realiza una evaluación del aporte de cada integrante en la solución propuesta para el proyecto.
- Se entrega los datos de prueba compartidos, en formato csv, con la etiqueta asignada por su modelo.
- La fecha máxima de entrega es el **domingo 15 de octubre, 22:00 pm.**

Bono

Los grupos que obtengan los mejores dos resultados de acuerdo con las métricas de desempeño de los modelos propuestos (0.5/5.0)



Etapas 2. Automatización de analítica de textos

Objetivos

- Automatizar un proceso replicable para aplicar la metodología de analítica de textos en la construcción de modelos analíticos.
- Desarrollar una aplicación que utilice un modelo analítico basado en aprendizaje automático y sea de interés para una organización, empresa o institución y en particular para un rol existente en alguna de ellas.
- Interactuar con un grupo interdisciplinario para validar y mejorar la calidad de la solución analítica planteada y del producto de software construido.

Contexto

El desarrollo de soluciones analíticas basadas en modelos analíticos para una organización incluye la participación de diferentes roles. En particular, en ese proceso de desarrollo,

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

haremos énfasis en los roles del científico de datos, ingeniero de datos e ingeniero de software.



El científico de datos tiene la responsabilidad de seguir un proceso replicable y de calidad, para entender los datos, prepararlos, construir el modelo analítico y evaluarlo, tanto a nivel cuantitativo como cualitativo. Este proceso fue realizado en gran medida en la etapa 1 de este proyecto. Por su parte, el ingeniero de datos tiene la responsabilidad de tomar el resultado entregado por el científico de datos, completarlo incluyendo elementos como pruebas, organización de código, validación de atributos de calidad (e.g., eficiencia), además de persistir el modelo generado en un repositorio específico para este fin. Adicionalmente, deben extender el código para incluir el registro o log, que permita en caso de falla, revisar las entradas al modelo, los datos utilizados para el entrenamiento y prueba del modelo y el resultado de la preparación. Finalmente, el ingeniero de software toma ese resultado como entrada de la aplicación web o móvil que debe diseñar y construir, la cual es utilizada por el usuario final para ingresar su solicitud al modelo, interactuar con el resultado y apoyar sus acciones o decisiones. Esta etapa 2 del proyecto se centra en este proceso de desarrollo de la solución analítica.

Descripción de la entrega

La entrega de esta etapa del proyecto contiene un documento que describe el proceso de despliegue de la solución analítica en el ambiente de producción de una organización. Esto incluye la automatización del proceso desde la preparación de los datos hasta la persistencia del modelo analítico construido, al igual que el dejar disponible una aplicación que toma el resultado del modelo y lo pone a disposición de un usuario final. Como parte de la entrega, adicionalmente, se debe incluir un video y todo el software desarrollado para el proyecto. En particular debe entregar:

(20%) Proceso de automatización del proceso de preparación de datos, construcción del modelo, persistencia del modelo y acceso por medio de API: Descripción del proceso e implementación realizados por el ingeniero de datos, acompañado del código. Debe utilizar *pipelines* para este proceso y un Framework para desarrollo de la API.

(40%) Desarrollo de la aplicación y justificación. Descripción del usuario/rol de la organización que va a utilizar la aplicación, la conexión entre esa aplicación y el proceso de negocio que va a apoyar (si aplica), y la importancia que tiene para ese rol la existencia de

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

esta aplicación. En este punto si lo requiere, ajuste la tabla de tipo de actores, entregada en la etapa 1 del proyecto.



Adicionalmente, el desarrollo de una aplicación web o móvil para interactuar con el resultado del modelo (y su calidad) a partir de un texto o textos datos por el usuario.

En este punto debe contar con el apoyo **del grupo de estadística** con el que está trabajando y mostrar evidencia de las mejoras que logró gracias a esa interacción. En este punto se espera que puedan describir las opciones que tuvieron al momento de definir la aplicación a desarrollar y el usuario-cliente, al igual, que la justificación, resultado del trabajo transdisciplinar. En esta descripción deben detallar, adicionalmente, el aporte dado por la persona de estadística que los acompañó.



(18%) **Resultados.** Video de máximo 5 minutos con la descripción y visualización en la aplicación de los resultados del modelo, que permita a un rol dentro de la organización comprenderlos y usarlos. El video publicado en el padlet respectivo (ver condiciones de entrega), debe simular la interacción del usuario final con la aplicación y describir dos acciones que puede realizar como resultado de dicha interacción, haciendo énfasis en la forma como el resultado del modelo aporta en esas acciones. En este punto debe contar con el apoyo del grupo de estadística con el que trabajó y mostrar evidencia de las mejoras que logró gracias a esa interacción. Lo más importante con respecto a esa interacción es revisar la validez de los resultados obtenidos con los modelos analíticos y garantizar que el uso que le están dando es correcto. En este punto, pueden mostrarle al estudiante del curso de estadística la aplicación final y recibir retroalimentación de él, realizando el rol de un posible usuario final. Como idea para este punto, les contamos que es posible vincular al estudiante de estadística en el desarrollo del video.

(10%) **Trabajo en equipo.**

- En este punto es importante comprender que su grupo está formado por estudiantes del curso de inteligencia de negocios y por estudiantes del curso de estadística. Con esta interacción se busca:
 - Trabajar de forma transdisciplinar para fortalecer sus competencias de argumentación y mejorar la calidad de sus trabajos, gracias al análisis desde diferentes puntos de vista

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

- Validar los resultados de los modelos analíticos construidos a la luz del vocabulario utilizado y las conclusiones que se dan con el fin de que tengan una argumentación formal y correcta.
- Para el trabajo en equipo se sugiere los siguientes roles para cada integrante. Un integrante puede desempeñar más de un rol.
 - **Líder de proyecto:** Está a cargo de la gestión del proyecto. Define las fechas de reuniones, pre-entregables del grupo y verifica las asignaciones de tareas para que la carga sea equitativa. Se encarga de subir la entrega del grupo. Si no hay consenso sobre algunas decisiones, tiene la última palabra.
 - **Ingeniero de datos:** Es responsable de velar por la calidad del proceso de automatización relacionado con la construcción del modelo analítico.
 - **Ingeniero de software responsable del diseño de la aplicación y resultados:** Se encarga de liderar el diseño de la aplicación y de la generación del video con los resultados obtenidos.
 - **Ingeniero de software responsable de desarrollar la aplicación final:** Se encarga de gestionar el proceso de construcción de la aplicación.
- Se recomiendan al menos las siguientes reuniones de grupo:
 - **Reunión de lanzamiento y planeación:** Para definir roles y forma de trabajo del grupo. Se genera lluvia de ideas sobre la forma de resolver el proyecto. Se sugiere en esta reunión integrar al grupo de estadística que le fue asignado para contactar si aún no lo han hecho, lo más pronto posible al estudiante de estadística ya que si lo dejan para última hora no hay garantía de que puedan lograr los objetivos. Determinar como grupo extendido (ustedes como estudiantes de BI y el estudiante de estadística) la forma como van a trabajar, los momentos en que se van a reunir, el canal a utilizar y la responsabilidad de cada uno. En particular revise con él la organización/empresa/institución y el rol dentro de ella, que se beneficia de la solución analítica que van a desarrollar, al igual que revisar y si se requiere ajustar la tabla de tipos de actores.
 - **Reuniones de seguimiento:** Se recomienda mínimo dos reuniones de seguimiento semanal cortas. También pueden ser correos de avance según lo defina el grupo. Pueden tener un tablero de control de las tareas, utilizando herramientas como Trello.
 - **Reunión de finalización:** Para consolidar el trabajo final, verificar el trabajo del grupo y analizar los puntos a mejorar para el siguiente proyecto. En esta reunión,

	Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20	
---	---	---

deben interactuar con el grupo de estadística para validar que se logra lo que definieron inicialmente.

- Describir los roles y las tareas realizadas por cada integrante del grupo, los tiempos, en número de horas, dedicados para realizar lo que le fue asignado, al igual que los retos enfrentados en el proyecto y las formas planteadas para resolverlos. Adicionalmente, de acuerdo con la planeación establecida para realizar el proyecto y la realidad en el desarrollo del mismo, reflexionen sobre la forma como repartiría 100 puntos entre los integrantes del grupo e identifique puntos a mejorar para la siguiente entrega del proyecto.



(12%) Sustentación y evaluación del aporte individual

Adicionalmente deben:

1. Entregar el conjunto de datos resultado de la fase de entendimiento y preparación de los datos. Requerido para realizar la sustentación.
2. El video construido
3. Especificar en el enlace de la wiki en el github de su proyecto (puede ser un repositorio público o compartirlo a través de githubfront). Todo el código debe quedar en el repositorio del proyecto del curso y referenciado en los entregables. Recuerden que pueden acceder a beneficios de estudiante en <https://education.github.com/pack>.

Evaluación y entregables

- El proyecto se realiza en grupos de 3 estudiantes. Los estudiantes pueden estar en distintas secciones.
- El documento a entregar tiene máximo 10 páginas (sin incluir portada, tabla de contenido, ni referencias), a una columna y con letra arial, tamaño 12.
- Deben publicar los resultados del proyecto usando el video generado, en el padlet <https://uniandes.padlet.org/mavillam/exposici-n-proyecto-anal-tica-de-texto-de-bi-202320-6s23hgij5y5tb29h> con clave: ProyBI232
- Deben entregar todo lo relacionado con el proyecto para poder ejecutarlo el día de la sustentación.

 <p>Universidad de los Andes Colombia</p>	<p>Ingeniería de Sistemas y Computación Pregrado ISIS-3301 – Inteligencia de Negocios Primer Proyecto Semestre: 2023-20</p>	
--	---	---

- Deben entregar el video realizado
- Se sustentará el proyecto y se realizará una evaluación del aporte de cada integrante en la solución propuesta para el proyecto.
- La fecha máxima de entrega es el **domingo 29 de octubre, 22:00 pm.**

ⁱ <https://www.un.org/sustainabledevelopment/es/development-agenda/>

ⁱⁱ <https://www.un.org/sustainabledevelopment/es/sustainable-development-goals/>

ⁱⁱⁱ <https://colombia.unfpa.org/es/unfpa-en-colombia>