



Universidad Nacional Autónoma de México
Escuela Nacional de Estudios Superiores
Unidad Morelia



Proyecto Final
Predicción del Crecimiento Significativo en Plantas

PRESENTA:

Alexis Uriel Aguilar Uribe

PROFESORES:

Dra. Marisol Flores Garrido

Dr. Luis Miguel García Velázquez

GRADO

Licenciatura en Tecnologías para la Información en Ciencias

Número de Cuenta: 424060075

Asignatura: Sistemas basados en conocimiento [Machine Learning]

A: 27 de Mayo del 2025

Índice

1. Introducción	2
2. Descripción de los Datos	3
3. Análisis Exploratorio de Datos	5
4. Metodología del Proyecto	6
5. Experimentos y Discusión de Resultados	7
6. Análisis de los Resultados	8
7. Conclusiones	9
Referencias Bibliográficas	10

1. Introducción

En la agricultura, como cualquier otra industria, se vuelve relevante la optimización de los recursos y ganancias, es decir, reducir los insumos consumidos mientras se incrementa la producción (tanto en calidad como en cantidad); todo lo anterior se traduce en aplicar mejoras en diferentes áreas y aspectos que convergen y se relacionan para generar ganancias y reducir costos en la agricultura. Para el caso de este proyecto, el interés se encuentra en el crecimiento de las plantas, bajo qué factores ambientales y de cuidado propician un crecimiento significativo en las plantas.

Para lograr el último punto, se tiene como objetivo el crear un modelo de aprendizaje supervisado para la clasificación del crecimiento significativo en base a los factores y mediciones relacionadas a su cuidado y ambiente.

2. Descripción de los Datos

El conjunto de datos que se emplearán para el proyecto se encuentra disponibles en [1], que es un conjunto de datos publicados en Kaggle por la propia comunidad. Se cuenta con siete columnas, donde seis de ellas son atributos y la otra el target, referenciando a la fuente del conjunto de datos, se tienen los siguientes atributos junto con su descripción y tipo de dato:

- **Soil_Type** [*String*]: El tipo o composición del suelo en el que las plantas están creciendo o se plantan.
- **Sunlight_Hours** [*Float*]: La duración o intensidad de la luz solar que las plantas reciben.
- **Water_Frequency** [*String*]: Qué tan seguido se riegan las plantas, se indica la frecuencia del riego.
- **Fertilizer_Type** [*String*]: El tipo de fertilizante usado para nutrir a las plantas.
- **Temperature** [*Float*]: Las condiciones de la temperatura ambiental bajo las cuales las plantas están creciendo.
- **Humidity** [*Float*]: El nivel de humedad en el ambiente alrededor de las plantas.
- **Growth_Milestone** [*Integer, Target*]: Descripción o marcadores que indican la etapa o eventos significativos en el proceso de crecimiento de las plantas.

Por último, el conjunto de datos consta de 193 instancias (filas), las diferentes instancias lucen de la siguiente manera:

Soil_Type	Sunlight_Hours	Water_Frequency	Fertilizer_Type
sandy	9.228	daily	none
sandy	9.774	weekly	chemical
clay	7.392	bi-weekly	none
clay	6.462	bi-weekly	organic
clay	8.846	weekly	organic
loam	5.985	bi-weekly	chemical

Temperature	Humidity	Growth_Milestone
33.804	32.815	0
32.549	61.377	1
31.100	68.600	0
27.517	34.175	1
27.700	56.800	1
29.757	57.476	0

3. Análisis Exploratorio de Datos

4. Metodología del Proyecto

5. Experimentos y Discusión de Resultados

6. Análisis de los Resultados

7. Conclusiones

Referencias Bibliográficas

- [1] gorororororo23, *Plant Growth Data Classification*, <https://www.kaggle.com/datasets/gorororororo23/plant-growth-data-classification/data>, 2024.