

# Proyecto Final – Aprendizaje Automático 2025

## COVID19 - Predicción de Nuevas Dosis de Vacunas Administradas en Argentina

Alumno: Gabriela Velasquez

Profesor: Mirabete

---

### 1. Introducción y Objetivo del Proyecto

El objetivo de este proyecto es construir un modelo predictivo que permita estimar la cantidad diaria de nuevas dosis de vacunas administradas en Argentina. Esta estimación se realiza en función de variables climáticas, estacionales y de movilidad poblacional.

El problema se aborda como una tarea de regresión supervisada, siendo la variable objetivo `new_vaccine_doses_administered`. La motivación principal es comprender los patrones que afectan la vacunación y prever su comportamiento para mejorar la planificación de recursos sanitarios.

### 2. Modelos aplicados:

- a. Regresión Linea
- b. Random Forest Regressor
- c. Support Vector Regressor (SVR)

### 3. Descripción del Dataset y Origen

El dataset utilizado proviene de registros diarios de vacunación junto con datos de movilidad y clima. Incluye las siguientes variables:

- `average_temperature_celsius`
- `rainfall_mm`
- `mobility_workplaces`
- `mobility_residential`
- `relative_humidity`
- Mes, Año
- `new_vaccine_doses_administered` (objetivo)

Periodo aproximado de recolección: 2021-2023

Cantidad de registros: 500

## Proyecto Final – Aprendizaje Automático 2025

### 4. Procesamiento ETL

- Carga de datos desde archivo CSV
- Limpieza de datos: imputación de valores nulos con cero
- División en variables predictoras y variable objetivo
- Estandarización de variables numéricas con StandardScaler

### 5. Análisis Exploratorio de Datos (EDA)

- Descripción estadística de todas las variables
- Visualización de distribuciones y tendencias
- Identificación de correlaciones entre variables

### 6. Preprocesamiento para el Modelado

- Escalado estandarizado
- División entre entrenamiento y prueba (80%-20%)

### 7. Modelado y Evaluación de Algoritmos

- Random Forest: MAE: 63.57, MSE: 21076.29,  $R^2$ : 0.73
- Regresión Lineal: MAE: 179.42, MSE: 66092.56,  $R^2$ : 0.15
- SVR: MAE: 130.66, MSE: 88513.21,  $R^2$ : -0.14

### 8. Métricas usadas:

- MAE, MSE,  $R^2$

### 9. Predicciones Futuras

Se generaron predicciones para los meses 6 a 12 de 2025, utilizando el modelo SVR y valores estimados para las variables independientes.

### 10. Conclusiones Finales

- Random Forest es el modelo más adecuado para este problema
- SVR se utilizó para completar el año en la visualización
- Las predicciones muestran un incremento estacional en invierno

## Proyecto Final – Aprendizaje Automático 2025

El modelo SVR fue utilizado para proyectar la cantidad de nuevas dosis de vacunas administradas durante los meses de junio a diciembre del año 2025. Para esto se utilizaron valores estimados de variables como temperatura promedio, precipitaciones, movilidad y humedad.

El gráfico de barras muestra un incremento progresivo de dosis predichas desde junio, alcanzando un pico máximo en agosto, y descendiendo gradualmente hacia diciembre. Esta tendencia es coherente con patrones estacionales observados en campañas pasadas, donde el invierno impulsa refuerzos en vacunación.

La predicción sugiere:

- Un patrón estacional marcado (invierno con mayor actividad)
- Que la movilidad laboral y la reducción del aislamiento pueden estar asociadas con un mayor volumen de vacunación

Conclusión:

El modelo permite anticipar los meses de mayor demanda, lo cual puede ser útil para planificar campañas, compras de insumos o logística sanitaria. Aunque SVR no fue el mejor modelo global, resultó funcional para completar predicciones en el ciclo anual completo y visualizar la estacionalidad esperada.