RESUMERN y ABSTRACT

INTRODUCCION

REDES N

MODELO

IA

PARA QUE: PEREDECIR CONT. TIMEPO ATMOSFERICO

OBJETIVOS

CONCEPTOS TEORICOS

TFM PABLO PUEIDE ESTAR MUY BIEN

DIFERENCIA MODELO FISICO, IA

TECNICAS Y HERRAMENTAS

METODOLOGIA DESING SCIENCE, SCRUM

LENGUAJES PROGRAMACION

MOTOR

ASPECTOS RELEVANTES (PARTE MAS IMPORTANTE)

DISEÑO MODELO

PRUEBA MODELO

FRONTEND

CONCLUSION

REFERENCIAS

ANEXO TECNICO

PLAN PROEYCTO

REQUSITOS

DISEÑO

APLICAICON WEB

DISEÑO MODELO

…

CODIGO

MANUAL DE USUARIO

RESUMEN

La contaminación atmosférica es uno de los grandes problemas a los que el mundo contemporáneo tiene que hacer frente. Por ello, existe una creciente necesidad desde, principal pero no exclusivamente, las grandes ciudades de conocer la evolución de los niveles de calidad del aire y poder anticiparse en la toma de decisiones.

Para esta tarea, la inteligencia artificial es de gran utilidad en la creación de una estimación que intente mejorar las obtenida por medios tradicionales como modelos teóricos. Concretamente, técnicas de *Deep Learning* y sus aplicaciones como las redes neuronales, debido entre otras cosas a su capacidad de procesamiento en paralelo, son idóneas para este cometido.

En este trabajo vamos a obtener un conjunto de datos sobre la ciudad de Madrid para crear un modelo con el que predecir los valores de diferentes contaminantes a lo largo del tiempo integrándolo en una aplicación web para la visualización de los resultados.

Palabras clave: redes neuronales, contaminación atmosférica, Deep Learning.

ABSTRACT

Air pollution is one of the major problems that the contemporary world has to face. Therefore, there is a growing need from, mainly but not exclusively, large cities to know the evolution of air quality levels and to be able to anticipate in decision making.

For this task, artificial intelligence is very useful in the creation of an estimate that tries to improve those obtained by traditional means such as theoretical models. Specifically, Deep Learning techniques and their applications such as neural networks, due, among other things, to their parallel processing capacity, are ideal for this task.

In this work we will obtain a set of data about the city of Madrid to create a model to predict the values of different pollutants over time by integrating it into a web application for the visualization of the results.

Keywords: neural networks, air pollution, Deep Learning.

INTRODUCION

Contexto

Desde la Revolución Industrial, los niveles de contaminación han ido aumentando enormemente, pero no ha sido hasta hace pocos años cuando la sociedad ha comenzado a tomar verdadera conciencia del asunto y los gobiernos han empezado a implantar medidas, quizás viéndose obligados por los malos pronósticos que la comunidad científica auguraba sobre el futuro de nuestro planeta a bastante corto plazo.

Este aumento de la polución se ha reflejado tanto en la aparición de nuevos problemas como en el agravante de algunos existentes, en cuestiones tan diversas como el aumento de las complicaciones respiratorias, la desaparición de especias animales y vegetales, el aumento de la temperatura o la acidificación de los océanos.

La lucha contra el cambio climático, en cualquiera de sus distintas facetas, es uno de los grandes desafíos a los que el mundo va a tener que afrontar las próximas décadas. Concretamente, la contaminación ambiental, definida como “*la presencia de componentes nocivos (ya sean químicos, físicos o biológicos) en el medio ambiente (entorno natural y artificial), que supongan un perjuicio para los seres vivos que lo habitan, incluyendo a los seres humanos*” (**Cuidemos el planeta (2018). “Contaminantes"), conlleva un riesgo para la salud humana y para la del planeta en general.**

**Además de la concienciación individual necesaria para intentar frenar este proceso, las administraciones públicas deben aplicar normas para disminuir la concentración de sustancias contaminantes en la atmósfera. Esto conlleva un análisis previo de los datos disponibles (recogidos a través de estaciones, similar al proceso seguido con los datos meteorológicos) para garantizar que las decisiones tomadas son las correctas y el grado con el que se está actuando es el adecuado, tanto a corto como a largo plazo.**

**Estas mediciones que realizan distintos organismos a lo largo de todo el mundo también pueden ser utilizadas en la predicción de los valores futuros de los contaminantes. Desde hace varias décadas existen modelos teóricos que, basándose en la física, intentan simular el comportamiento de un sistema para obtener tendencias, pero usando la capacidad de procesamiento de los computadores actuales además de los avances en inteligencia artificial se busca mejorar estas predicciones.**

**En nuestro caso, aplicando técnicas de Machine Learning y Deep Learning usaremos los datos provistos por las agencias para crear modelos que permitan predecir con el mínimo error posible los niveles de contaminación de fechas venideras, previniendo episodios de alta contaminación y las consecuencias drásticas que estos conllevan.**

Contaminantes principales y límites

Existen gran cantidad de sustancias nocivas presentes en el aire, aunque la EPA (Agencia de Protección Ambiental de Estados Unidos [https://www.cdc.gov/air/pollutants.htm](https://www.cdc.gov/air/pollutants.html)) identifica seis contaminantes como los principales, regulados mediante valores límite basados en los efectos que provocan tanto para la salud pública como para el medio ambiente. Son el monóxido de carbono, el plomo, los óxidos de nitrógeno, el ozono troposférico, la materia particulada y los óxidos de azufre. Se pueden clasificar en primarios o secundarios dependiendo si son generados por procesos humanos/naturales o generados por reacciones químicas de los primarios.

Para este trabajo, vamos a recoger los datos de los que más influencia tienen en Madrid para posteriormente realizar la predicción de cada uno de ellos. El límite marcado para cada uno de ellos viene dado por la Comunidad de Madrid (<http://gestiona.madrid.org/azul_internet/html/web/2_3.htm?ESTADO_MENU=2_3>), pudiendo ser horarios/diarios (no pueden superarse más de un determinado número de veces al año) o anuales.

|  |  |  |
| --- | --- | --- |
| Contaminante | Tipo | Límite |
| PM2.5 | Primario y secundario | 25 µg/m3 (media anual) |
| PM10 | Primario y secundario | 40 µg/m3 (media anual) |
| O3 | Secundario | 120 µg/m3 (máxima diaria) |
| NO2 | Primario y secundario | 40 µg/m3 (media anual) |
| SO2 | Primario | 125 µg/m3 (media diaria) |

<https://www.comunidad.madrid/servicios/salud/calidad-aire-salud>

Materia particulada: uno de los indicadores más comunes. Son una mezcla de partículas, tanto sólidas como líquidas, suspendidas en el aire. Estudios realizados en la Unión Europea (<https://www.eea.europa.eu/es/themes/air/intro>) muestran que son el agente contaminante más nocivo para las personas, debido a la capacidad de penetrar en el cuerpo. Dependiendo del grosor, se pueden clasificar en varios tipos. Los que mediremos son:

PM2.5: las partículas con diámetro menor a 2,5 micras. Son las más dañinas para la salud debido a que pueden llegar al torrente sanguíneo, lo que provoca efectos en el sistema cardiovascular además del respiratorio. Están formadas por elementos más tóxicos procedentes principalmente del tráfico urbano, y son capaces de mantenerse en el aire por más tiempo que las de mayor tamaño.

PM10: las partículas con diámetro menor a 10 micras. Pueden ser inhaladas por el sistema respiratorio, aunque no atraviesan los alveolos pulmonares como las anteriormente mencionadas. El tiempo de permanencia es menor, de horas en vez de días. El 77,9% (<https://prtr-es.es/particulas-pm10,15673,11,2007.html>) de emisiones proceden del polvo resuspendido existente en la atmósfera.

O3: el ozono troposférico es aquel presente a nivel de suelo. Aunque el ozono es fundamental para proteger de la radiación UV en la estratosfera, en niveles más cercanos la tierra es muy perjudicial. Causa problemas de salud entre los que se encuentran dificultades respiratorias, daño a los pulmones y asma. También tiene efectos adversos sobre la vegetación, interfiriendo con el proceso de la fotosíntesis.

NO2: el principal contaminante entre los óxidos de nitrógeno. Producido por fábricas, vehículos y quema de residuos. Es un gas tóxico que además incrementa los niveles de PM2.5. Tiene gran influencia sobre enfermedades respiratorias y provoca la lluvia ácida, con severos efectos sobre la fauna y flora.

SO2: el principal contaminante entre los óxidos de azufre y el más peligroso. La mayor fuente de emisión son las fábricas industriales. También incrementan los niveles de materia particulado al igual que los óxidos de nitrógeno. Sobre las personas, afectan principalmente al sistema respiratorio. Sobre el medio ambiente, provocan una disminución en el crecimiento de plantas y árboles.

AQI (Air Quality Index)

Es un índice utilizado por agencias gubernamentales para facilitar el acceso a los datos de contaminación y marcar un criterio común, así no teniendo que utilizar las concentraciones reales de los contaminantes y sus respectivos límites. En la Unión Europea no se usa un índice, sino que cada país tiene su propio criterio para comunicar sus niveles, mientras que por ejemplo en Estados Unidos la Agencia de Protección del Medio Ambiente establece una escala de 0 a 500 con un código de colores. (<https://www.airnow.gov/aqi/aqi-basics/>)

|  |  |  |
| --- | --- | --- |
| Valor del índice de la calidad del aire | Amenaza para la salud | Colores |
| 0 a 50 | Buena | Verde |
| 51 a 100 | Moderada | Amarillo |
| 101 a 150 | Insalubre para grupos sensibles | Naranja |
| 151 a 200 | Insalubre | Rojo |
| 201 a 300 | Muy insalubre | Morado |
| 301 a 500 | Peligrosa | Granate |

En este trabajo vamos a utilizar valores del índice de la calidad del aire en lugar de valores absolutos, utilizando el AQI estadounidense, tanto en la recogida de datos pasados, actuales y en la propia predicción de los valores futuros.

Inteligencia artificial y redes neuronales

El modelo de nuestro sistema se crea utilizando redes neuronales. Son una de las más importantes técnicas dentro del mundo de la inteligencia artificial y supusieron una revolución en este campo, difundiéndose enormemente por su capacidad para dar buenos resultados en ámbitos diversos.

ALGUNA IMAGEN, EJ NEURAL NETWORKS TUTORIAL 1

A partir de unos datos de entrada, la red neuronal los procesa y se crea un modelo, de los que hay varios tipos. Los más comunes son los de aproximación, clasificación y predicción. En este caso, por razones obvias, tendremos un modelo de predicción, usados para pronosticar el estado futuro de un sistema a partir de observaciones pasadas sobre él mismo. Otros trabajos de predicción con redes neuronales incluyen pronósticos de ventas, macroeconómicos o de acciones de marketing.

Las características de la red neuronal influyen en el buen funcionamiento del modelo, siendo el mayor éxito posible la generalización a otros problemas. En nuestro caso, un ejemplo sería obtener buenos resultados en la ciudad de Barcelona a partir del modelo construido para Madrid. Para ello existen distintas topologías de redes neuronales, que pueden adecuarse más o menos a las características que buscamos, por lo que se entrará en detalle y experimentará con la finalidad de conseguir un error lo más pequeño posible en los valores alcanzados.

OBJETIVOS

El objetivo principal del trabajo es lograr obtener una predicción fiable del AQI para cada uno de los cinco contaminantes escogidos para la ciudad de Madrid.

Por lo tanto, es una parte fundamental minimizar el error del modelo para conseguir la mayor precisión posible en los valores de salida. Para ello, es necesario encontrar la red neuronal adecuada, ajustando correctamente los parámetros necesarios y utilizando los algoritmos existentes convenientemente.

Para esta misión, es imprescindible extraer los datos históricos necesarios con exactitud desde fuentes fiables, ya que se utilizarán en el entrenamiento de la red neuronal y las posteriores pruebas. Además de ellos, los datos en tiempo real también son necesarios para poder adaptar el modelo lo máximo posible y que así no haya demora en el sistema.

Un objetivo secundario de este trabajo es permitir a un usuario sin conocimientos expertos en el tema poder visualizar los valores obtenidos con el modelo de manera sencilla e intu

CONCEPTOS TEORICOS

-Conjunto de datos

-Red neuronal

-Funcion de costo (error)

-Algortimos de entrenamiento