Global Warming: Electric Cars as a Potential Remedy

Problem definition

With the rise in air pollution in recent decades, the Earth's average temperature has gradually increased, and this phenomenon is described as global warming [1]. With continuous research and development focused on global warming and pollution, scientists and researchers worldwide are working towards solving this problem. The importance of keeping global warming at bay and reducing air pollution has long been recognised internationally, with the Paris agreement that aims to keep average global warming below 2 degrees Celsius [2], efforts are ongoing in all the participating countries to create a better future for our planet.

Global warming affects all species, with disastrous effects being observed worldwide. Plants, animals, communities, and ecosystems are affected by even a slight rise in the temperature of our Earth. The consequences are the extinction of plant species at lower latitudes [3] extinction of animal species [4]. The current repercussions of global warming are the rise in natural disasters such as flash floods, storms, wildfires, droughts, and heat waves. The indirect effects include the reduction of biodiversity and habitable areas, leading to the movement and degradation of community, public health, and socioeconomic conditions in most world countries [5].

This research project aims to find more insight into global warming and its correlation to carbon dioxide from different countries. This project intends to contribute its insights towards the betterment of humanity, scientifically, and provide more information as to what path we can pursue toward a better future. The fields that could benefit from the execution of this project are Government, Law and Public Policy, Science and Technology, and Education.

Background

A preliminary literature review shows that much research has been done to study the effects of pollution and global warming. The primary source of air pollution is road transportation, which contributes to the build-up of greenhouse gases comprising millions of tonnes of CO2 emissions [6][7]. Looking deeper into the data it can be observed that the entire transportation industry accounts for 24% of global carbon dioxide emissions, with road transportation responsible for 72% of transport-related emissions [8]. This can be linked to the rapid industrialisation and globalisation of the world, with developed and developing countries aiming for economic prosperity [15].

Data sources reveal that the USA is the top polluter, followed by the UK and China in third place with a steep rise in emissions [9]. Electric vehicles have gained popularity due to advanced technological development and declining costs. Environmental awareness has promoted usage with growth in EV adoption [10]. Electric cars solely powered by batteries benefit the environment, as they reduce air pollution [11]. The advantages of EV adoption have been discussed extensively in the literature such as, [16] where we can observe that the life-cycle emissions of the electric vehicles present in the market today are much lower than the life-cycle emissions of the most fuel-efficient fossil-fuel-based vehicles.

According to the researchers Sebestyen, Czvetko, and Abonyi in their paper only adding data regarding our climate is not sufficient to find a perfect solution for climate change, including other social environmental factors, will generalize our model much more. To solve scientific issues, data science is one of the most common tools. As mainstream business models have shown great support towards

this cause, the complexity of the issues has also increased tremendously due to that we need to include our own analysis methods regarding our data immediately. Although various fields can benefit from the results of other fields, new methods are necessary. The current course of climate and sustainability studies is propelling models to be much more complex, large, and of a higher resolution which demands a multifaceted approach to climate studies [17].

Proposed method

Data

The datasets that are being utilised to carry out this research project are "Co2_Emissions", "Global_Temperatures" and "Electric_Cars". All three datasets were provided by the Middlesex University Department of Computer Science for research. The "Co2_Emissions" dataset has nine features, of which two are categorical, and the other seven are quantitative, with 59,621 observations for each feature. The attribute 'country' represents the countries from which the data has been taken, and 'country code' represents the countries in abbreviated form. 'Calling code' has values of the telephonic calling codes of each country, and 'Year' represents the period during which the data was collected. 'CO2 emissions' contains the values of CO2 emitted by that specific country in the given year and are one of the main attributes that will be used from the dataset. 'Population' represents the population of the country during that specific year. Finally, 'Area', '% of World', and 'Density' are geographical qualities of the data representing the land area, what percentage of the Earth's total land area it constitutes, and the country's population density, respectively. All the attributes of this dataset will be used for our research.

Moving to our following dataset, the "Global_Temperature" dataset has more attributes than the other two datasets used in this project, with 30 attributes and 2094 rows of observations. This dataset has observations of the temperature at different regions of the Earth for every month from 1978 to 2022. The attribute 'Year' and 'Mo' represents the year and month. After which, we can observe triads of data with temperature data of the land and ocean and names of the different hemispheres and sub-regions of the Earth. 'USA48' is used to indicate data about the temperature of the 48 states in the US, and 'USA49' is the same as 'USA48' but includes data about Alaska, 'AUST' is the temperature data of Australia and 'Troposphere' is indicative of which region of the tropospheric layer the data is obtained from. None of the attributes will be omitted in our research.

The final data set is the "Electrics Cars" dataset. It has six attributes, and all of them are quantitative, with 14 rows of observations. The attributes are 'Year' which indicates the year, 'BEV Average Price' which indicates the average price of an electric vehicle with battery, 'Global sales volume' which represents the global sales of electric cars, 'Mileage' represents the average mileage of the electric vehicles per full charge in that year, 'Lithium Ion Battery Price USD' represents the average price of the electric vehicle's battery, 'Average price of new car' represents the average price of a new electric car in that year. All the attributes of this dataset will be used for our research. From our initial observation of the data, we can observe that all the data sets are balanced, and the issue of imbalanced data has not been observed.

Research questions / Hypotheses

The proposed hypothesis for this project is ‘Will the increased usage of electric vehicles aid in decreasing CO^2 emissions, therefore leading to reduced global warming?’.

Feature Engineering

Feature engineering is carried out in the "Global_Temperatures" dataset, the triads of data described earlier are highly correlated to one another, hence we propose to reduce it from three features to one by performing dimensionality reduction on each pair. The features Northern Hemisphere and Northern Extremes have a high correlation therefore, we can reduce it from six columns to one. The same can be done with Southern Hemisphere and Southern Extremes. After performing dimensionality reduction, the shape of the dataset became (2094,11) with minimal loss of information while maintaining 99.5% of the original variance. A new feature has been created while feature engineering; it is what the population of each country owns an electric car. This has been done by combining "Electric car prices" and the "CO2 emissions" dataset. We combine them in the column 'Year' and for the years in which we do not have any data about electric cars, we substituted it with the value zero.

According to a statistical study conducted by BloombergNEF, it has been found that in 2022 around 13% of the cars sold will be electric on a global scale. We can use the data that in each country, around 3% of the population owns an electric car [13,14]. So, we take 3% of each country's population and derive the number of electric cars owned by that country which will help us in analysing whether electric vehicles can aid in decreasing CO2 emissions in the long run. Next, we categorise each country according to the regions and subregions of the Earth they are situated in so that we can combine this data with the "Global_Temperatures" dataset, which will help in modelling the data.

Data processing

For Data processing, we link all three datasets using the common attribute 'Year' then, in the "Co2_Emissions" dataset we convert the density value into a single quantitative value by omitting /km^2 for the % feature, we convert the observations into values ranging from 0 – 1. E.g., 2%->0.02. For the countries and their codes, we use their calling codes as they are an encoded version of their names, hence we can omit the attributes ‘country’ and ‘country code’. In the "Global_Temperatures" dataset we can observe that many columns display a high degree of skewness; for this project, we have set a threshold of 1. We must perform standard scaling on all columns having values of temperatures as we need to perform PCA on them, this will remove skewness. In the “ElectricCarPrices” dataset, the missing value of Global Sales of BEV has been populated with the value '913,479' [14]. The data type of the feature 'Year' has been changed to the date-time datatype from the object data type, and all the other columns have been converted from int to float.

Coming to missing values in our data, the "Global Temperatures" dataset has 11 missing values, 9 of them are due to unavailability of data as it was not collected during that time, hence they are omitted; 2 of them are missing values for ‘AUST’ which has been populated with the mean of the values preceding and proceeding the missing value. The categorical feature troposphere will be encoded with the OneHotEncoding method. Now we have combined all three datasets into one. Cleaning and removal of null values have also been done.

Modelling

Our target feature will be CO2 Emissions, and our input datasets will be the combined data set that we created while data processing, on which we will be performing Regression. After that, if we input new data about 2023, i.e., the number of electric cars owned and the temperature details, we should observe a reduction in CO2 Emissions for that country from its previous years; this would mean that we cannot reject our hypothesis.

Bibliography

[1] John Houghton 2005 Rep. Prog. Phys. 68 1343

[2] https://unfccc.int/sites/default/files/english_paris_agreement.pdf , UNITED NATIONS (2015) Adoption of the Paris Agreement

[3] Hughes, L. (2000). Biological consequences of global warming: is the signal already apparent? In Trends in Ecology & Evolution (Vol. 15, Issue 2, pp. 56–61). Elsevier BV. https://doi.org/10.1016/s0169-5347(99)01764-4

[4] https://www.wwf.org.uk/sites/default/files/2018-03/WWF_Wildlife_in_a_Warming_World.pdf WWF's (2018) Wildlife In a Warming World

[5] Watts, N., Amann, M., Arnell, N., Ayeb-Karlsson, S., Belesova, K., Boykoff, M., Byass, P., Cai, W., Campbell-Lendrum, D., Capstick, S., Chambers, J., Dalin, C., Daly, M., Dasani, N., Davies, M., Drummond, P., Dubrow, R., Ebi, K. L., Eckelman, M., … Montgomery, H. (2019). The 2019 report of The Lancet Countdown on health and climate change: ensuring that the health of a child born today is not defined by a changing climate. In The Lancet (Vol. 394, Issue 10211, pp. 1836–1878). Elsevier BV. https://doi.org/10.1016/s0140-6736(19)32596-6

[6] *(J. Hu, R. Wood, A. Tukker, and B. Boonman H.and de Boer. 2019. Global transport emissions in the Swedish carbon footprint. Journal of Cleaner Production, 226, 210–220.)*

[7] https://www.iass-potsdam.de/en/output/dossiers/air-pollution-and-climate-change Institute for Advanced Sustainability Studies (IASS 2022) Air Pollution and Climate Change

[8] *S. Wang and G. Mengpin. 2019. Everything you need to know about the fastest-growing source of global emissions: transport.*

[9] Doris Klingelhöfer, Ruth Müller, Markus Braun, Dörthe Brüggmann and David A. Groneberg Climate change: Does international research fulfil global demands and necessities?

[10] Matteo Muratori et al 2021 Prog. Energy 3 022002 The rise of electric vehicles—2020 status and future expectations

[11] Lie, Tek Tjing & Prasad, Krishnamachar & Ding, Ning. (2017). The electric vehicle: a review. International Journal of Electric and Hybrid Vehicles. 9. 49. 10.1504/IJEHV.2017.10003709.

[12] https://policyadvice.net/insurance/insights/electric-car-statistics/ Alex Kopestinsky (2022) Electric Car Statistics In The US And Abroad

[13]]https://www.iea.org/data-and-statistics/data-product/global-ev-outlook-2022 IEA Publications (2022) Global EV Outlook 2022

[14] https://cleantechnica.com/2022/08/01/record-ev-sales-month-globally-12-share-for-bevs/ Jose Pontes (2022) Record EV Sales Month Globally! 12% Share For BEVs

[15] Peter Newell, Matthew Paterson, Climate Capitalism: Global Warming and the Transformation of the Global Economy

[16] Knobloch, F., Hanssen, S., Lam, A., Pollitt, H., Salas, P., Chewpreecha, U., Huijbregts, A. J., & Mercure, F. (2020). Net emission reductions from electric cars and heat pumps in 59 world regions over time. *Nature sustainability*, *3*(6), 437. https://doi.org/10.1038/s41893-020-0488-7

[17] Sebestyén, V., Czvetkó, T., & Abonyi, J. (2021). The Applicability of Big Data in Climate Change Research: The Importance of System of Systems Thinking. *Frontiers in Environmental Science*. https://doi.org/10.3389/fenvs.2021.619092