CO_2 WEB APP PREDICTION SYSTEM

By: Folasewa Abdulsalam, Adesanya Adenike, Ibrahim Murtala, Amaobichukwu Chibuzor

ABSTRACT.

They say too much of a good thing is a bad thing, this applies to atmospheric CO_2: In higher concentrations, it is a damaging pollutant. Hence, industries that contribute to higher CO_2 emission due to their production activities need to be able to use supervised machine learning techniques to predict CO_2 emission and unsupervised machine learning techniques to discover new materials with low CO_2 emission rate in their production processes. The case study for this project is the cement production industry. Carbon dioxide is an important component of the atmosphere because it plays multiple roles in keeping Earth's climate stable. CO_2 and other greenhouse gases are an essential part of the recipe because they trap heat in the atmosphere. With no CO_2, Planet Earth would be in a perpetual ice age; However too much of it overheats the planet.

The deployed model is deployed as a web application that is capable of predicting CO_2 emission from the masses of cement materials. A dataset containing cement materials is established for training and validating the model. The proposed model achieved an accuracy of 97%.

Below is the snapshot of the dataset used in the model building.

	Limestone	Shale	Iron ore	CaO	SiO2	Al2O3	Fe2O3	LSF	AM	SM	Emission factor	Co2
0	33.10000	5.784	1.116	25.60	8.800000	2.00	1.200	92.2	1.7	2.75	0.502400	20.096000
1	165.50000	28.920	5.580	129.00	43.600000	11.00	5.000	93.3	2.2	2.73	0.506300	101.260000
2	859.46148	27.500	5.300	121.20	230.621364	8.60	5.320	84.9	1.6	3.30	0.500830	521.746125
3	278.73000	46.000	9.234	224.01	77.980000	11.60	9.580	94.0	1.2	3.70	0.514175	175.847850
4	216.79000	35.777	6.384	179.55	56.924000	17.29	6.118	97.7	2.8	2.43	0.502359	140.946750

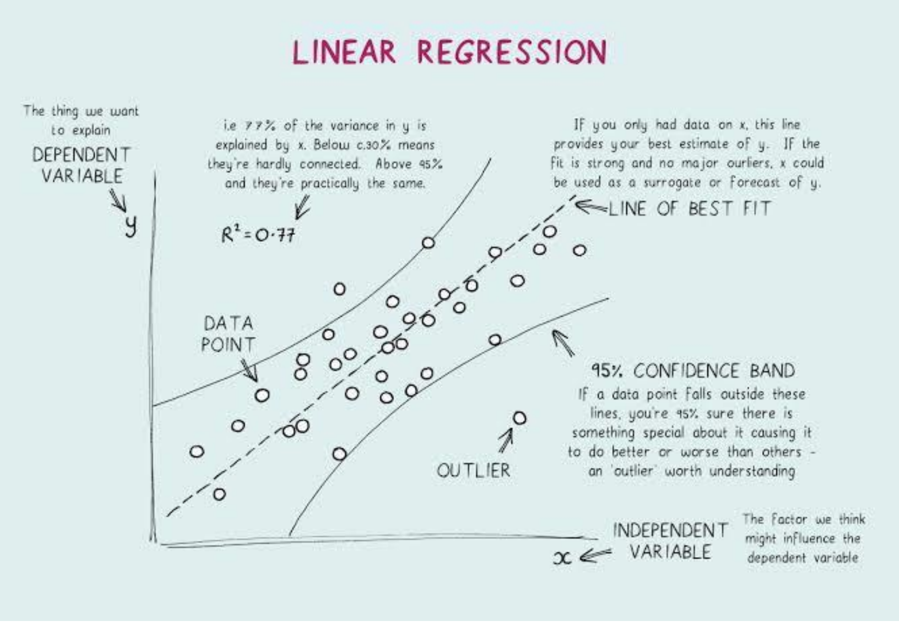
INTRODUCTION

Cement manufacturing industries contribute to greenhouse gases both directly through the production of carbon dioxide. In the process of making clinker which is the key constituent of cement that emits the largest amount of CO_2 in cement making.

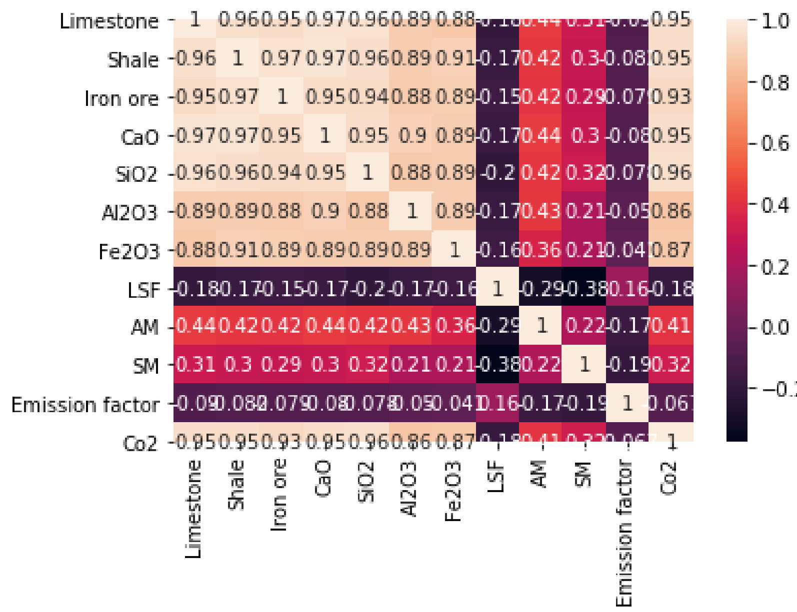
Being a significant CO_2 emitter, measures have to be considered to reduce the CO_2 emissions. There have been existing technologies for CO_2 reduction, and this prompted for the development of a prediction system that predicts CO_2 emission from clinker production materials. The solution proposed is a deployed web app that gives room for the user to input the masses of cement materials. The regression algorithm used is the linear regression.

METHODOLOGY

The method we are adopting to predict emission of CO_2 is supervised learning using Linear Regression.



To completely model this problem, a dataset containing 12 columns and 70 rows of data which was formulated by the problem stakeholders using engineering principles guiding the production of cement and data gotten from experiments. See figure below for the correlation among features.



DATA PREPROCESSING

> Missing data were filled with the mean of the data.

> Features with least correlation were dropped.

> Standard scaler was used in scaling down the features.

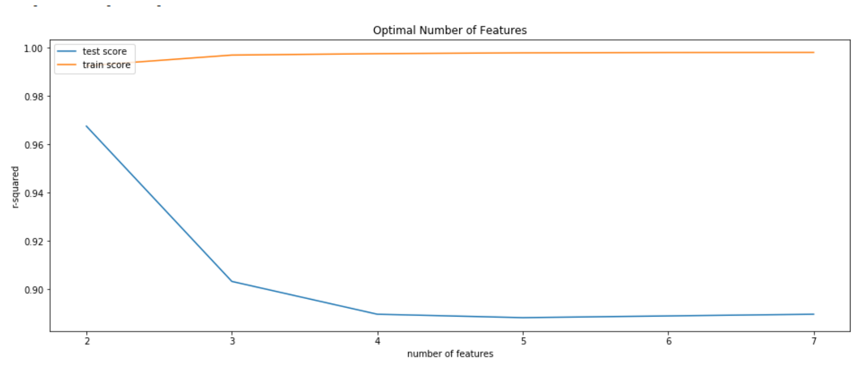
MODEL ARCHITECTURE

The architecture used in modelling this problem is a Linear Regression model. Grid Search Cross validation of k-folds=5 was used with an hyperparameter range between 2 and 8 and estimator = rfe. The estimator Recursive Feature Elimination (rfe) was used because it would select features based on how well the affect the model's performance. The diagram below shows the model architecture

	mean_fit_time	std_fit_time	mean_score_time	std_score_time	param_n_features_to_select	params	split0_test_score	split1_test_score	s
0	0.033492	0.021280	0.012882	0.015777	2	{'n_features_to_select': 2}	0.917891	0.941030	
1	0.018232	0.003198	0.003130	0.006259	3	{'n_features_to_select': 3}	0.739803	0.807324	
2	0.016933	0.002609	0.000000	0.000000	4	{'n_features_to_select': 4}	0.709101	0.774387	
3	0.011406	0.006083	0.008270	0.007049	5	{'n_features_to_select': 5}	0.708315	0.781287	
4	0.008225	0.003230	0.003028	0.003744	6	{'n_features_to_select': 6}	0.714509	0.778166	
5	0.004594	0.000488	0.002395	0.000487	7	{'n_features_to_select': 7}	0.714021	0.777215	

split2_test_score	split3_test_score	...	mean_test_score	std_test_score	rank_test_score	split0_train_score	split1_train_score	split2_train_score	split3_train_s
0.989239	0.993380	...	0.967457	0.031943	1	0.984383	0.998082	0.995241	0.99
0.987250	0.993734	...	0.903205	0.108007	2	0.999341	0.999558	0.995866	0.99
0.976604	0.995176	...	0.889628	0.122666	3	0.999871	0.999905	0.996408	0.99
0.970119	0.988687	...	0.888206	0.119584	6	0.999886	0.999906	0.996416	0.99
0.971083	0.989263	...	0.888921	0.118359	5	0.999887	0.999910	0.996644	0.99
0.975189	0.989819	...	0.889627	0.119408	4	0.999887	0.999910	0.996697	0.99

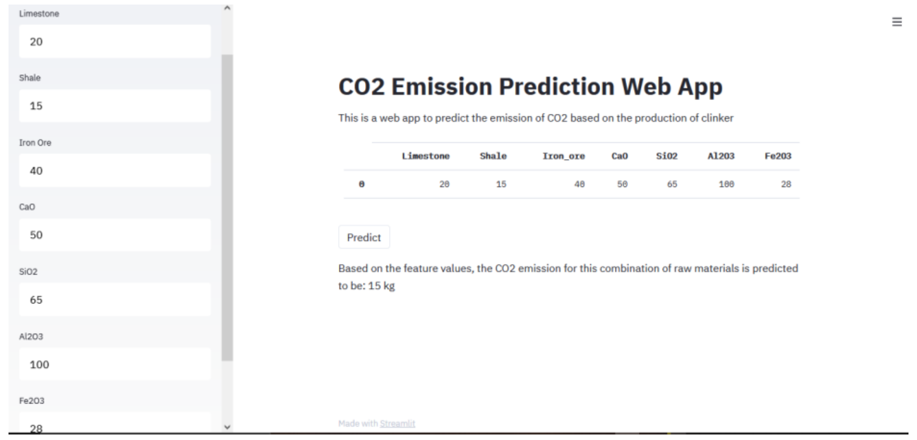
Plotting the CV results:



MODEL DEVELOPMENT

Streamlit was used in the model deployment and the figure below is a screenshot of the web app.

Link: https://drive.google.com/file/d/1k1lP7a-dl76tpCq0UxzyUddfd3zxcZOU/view?usp=sharing



RESULT

The following are the obtained results of the modelling. The metric used in evaluating this architecture is 'accuracy'.

REFERENCES

1. https://article.sapub.org/10.5923.j.control.20201001.03.html
2. https://www.understanding-cement.com/bogue.html
3. https://www.hindawi.com/journals/mpe/2012/392197/
4. https://www.researchgate.net/publication/269127824_A_Knowledge-Based_System_for_Mix_Design_of_Concrete_Containing_Pozzolanic_Materials
5. https://www.hindawi.com/journals/ace/2021/6682283/