Analyze_and_Predict_the_Effects_of_Automotive_Features_o... (3).pdf

ORIGINALITY REPORT

4%	3%	2%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	patents.justia.com Internet Source	<1%
2	Submitted to CSU, San Diego State University Student Paper	<1%
3	Submitted to London School of Business and Management Student Paper	<1%
4	123docz.net Internet Source	<1%
5	www.sciencepublishinggroup.com Internet Source	<1%
6	repository.sustech.edu Internet Source	<1%
7	H L Gururaj, Francesco Flammini, V Ravi Kumar, N S Prema. "Recent Trends in Healthcare Innovation", CRC Press, 2025 Publication	<1%
8	www.bartleby.com Internet Source	

		<1%
9	Submitted to University of Queensland Student Paper	<1%
10	Submitted to University of Hull Student Paper	<1%
11	biblio.ugent.be Internet Source	<1%
12	dspace.uiu.ac.bd Internet Source	<1%

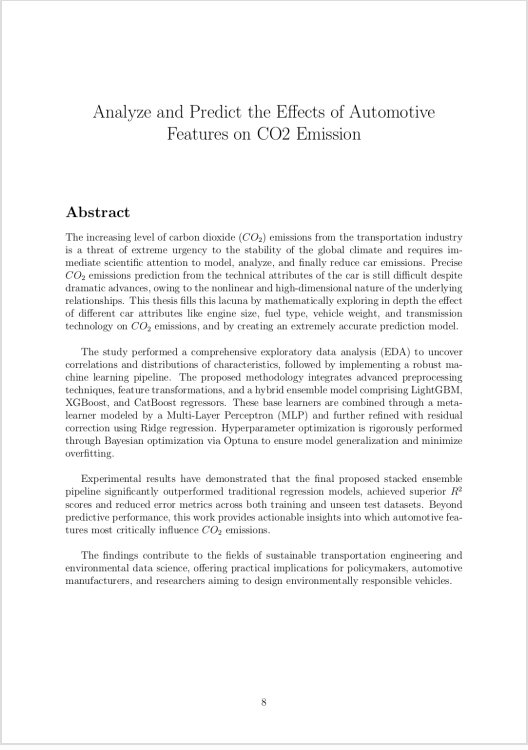
Exclude quotes	On	Exclude matches	< 14 words
Exclude bibliography	On		

Digital Receipt

This receipt acknowledges that Turnitin received your paper. Below you will find the receipt information regarding your submission.

The first page of your submissions is displayed below.

Submission author:	Shashvat Jain
Assignment title:	Thesis
Submission title:	Analyze_and_Predict_the_Effects_of_Automotive_Features_on_...
File name:	Analyze_and_Predict_the_Effects_of_Automotive_Features_on_...
File size:	1.55M
Page count:	32
Word count:	6,783
Character count:	44,125
Submission date:	02-May-2025 01:50PM (UTC+0530)
Submission ID:	2661927193



Analyze and Predict the Effects of Automotive Features on CO2 Emission

Abstract

The increasing level of carbon dioxide (CO_2) emissions from the transportation industry is a threat of extreme urgency to the stability of the global climate and requires immediate scientific attention to model, analyze, and finally reduce car emissions. Precise CO_2 emissions prediction from the technical attributes of the car is still difficult despite dramatic advances, owing to the nonlinear and high-dimensional nature of the underlying relationships. This thesis fills this lacuna by mathematically exploring in depth the effect of different car attributes like engine size, fuel type, vehicle weight, and transmission technology on CO_2 emissions, and by creating an extremely accurate prediction model.

The study performed a comprehensive exploratory data analysis (EDA) to uncover correlations and distributions of characteristics, followed by implementing a robust machine learning pipeline. The proposed methodology integrates advanced preprocessing techniques, feature transformations, and a hybrid ensemble model comprising LightGBM, XGBoost, and CatBoost regressors. These base learners are combined through a meta-learner modeled by a Multi-Layer Perceptron (MLP) and further refined with residual correction using Ridge regression. Hyperparameter optimization is rigorously performed through Bayesian optimization via Optuna to ensure model generalization and minimize overfitting.

Experimental results have demonstrated that the final proposed stacked ensemble pipeline significantly outperformed traditional regression models, achieved superior R^2 scores and reduced error metrics across both training and unseen test datasets. Beyond predictive performance, this work provides actionable insights into which automotive features most critically influence CO_2 emissions.

The findings contribute to the fields of sustainable transportation engineering and environmental data science, offering practical implications for policymakers, automotive manufacturers, and researchers aiming to design environmentally responsible vehicles.

8