

Predicting the Sales on Online Retail Products using Machine Learning Algorithm

PRESENTER:
Bhanu Prakash Reddy VENNAPUSA

AIM
The research mainly aims at predicting online sales of retail products using Machine Learning Algorithms, specifically Lasso Regression, Ridge Regression, and Neural Networks, including sentiment score.

METHODOLOGY
Quantitative research is well-suited for this study as it allows for the systematic analysis of numerical data and the application of statistical models to make predictions and draw conclusions.

MODEL BUILDING
Lasso Regression: It uses shrinkage and L1 regularization techniques to reduce variables to Zero.
Ridge Regression: It uses shrinkage and L2 regularization techniques to reduce variables near Zero.
Neural Network: Understands Unstructured and Complex datasets to make decisions and understands the nonlinear relationship between dependent and independent variables.

OPTIMISATIONS
To improve the model accuracy this study uses three optimisations such as Optuna, Particle Swarm(PSO) and Differential Evolution(DE) Optimisation.

Ridge Regression	Dataset D1	Optimisations	RMSE	MAE	TRAINING TIME(S)	TESTING TIME(S)
		Optuna	7.33e-09	5.43e-09	0.0007	0.001
		PSO	1.43e-13	1.09e-13	0.757	0.002
	Dataset D2	DE	1.66e-10	1.23e-10	0.004	0.0003
		Optuna	2.87e-06	2.19e-06	0.0005	0.0008
		PSO	4.49e-14	2.88e-14	0.217	0.0007
		DE	2.43e-10	1.85e-10	0.0019	0.0009

Optimisation comparison for Ridge Regression

Ridge Regression With Particle Swarm Optimisation (PSO) Proved as a best model for Sales Prediction.

	MODELS	RMSE	MAE	TRAINING TIME(S)	TESTING TIME(S)
DATASET D1	Lasso Regression	0.004	0.003	0.09	0.0003
	Ridge Regression	1.66e-08	1.23e-08	0.04	0.0003
	Neural Network	32.83	27.30	1.20	0.06
DATASET D2	Lasso Regression	0.0003	0.0002	0.04	0.0003
	Ridge Regression	2.42e-08	1.85e-08	0.11	0.0003
	Neural network	25.17	19.90	0.98	0.46

Sales demand prediction model output for D1 and D2



Take a picture for Project Code and Datasets

MODEL EVALUATION
For both datasets, the Ridge Regression model outperforms Lasso Regression and Neural Network in terms of metrics like RMSE, MAE, and prediction time, making it the best accurate model for sales prediction.

Models	ARIMA	SARIMA	LSTM
RMSE	8.681	7.274	3.917
ME	2.135	3.932	0.470
MPE	0.030	0.036	0.004
MAE	6.481	6.010	3.257

Sirisha, U. M., Belavagi, M. C. & Attigeri, G., 2022. Profit Prediction Using ARIMA, SARIMA and LSTM Models in Time Series Forecasting: A Comparison. IEEE Access, Volume 10, pp. 124715-124727.

The Ridge Regression model outperforms existing ARIMA, SARIMA, and LSTM models in achieving 1.66e-08 and 2.42e-08 RMSE values before optimisation.

CONCLUSION
Ridge Regression outperformed the Lasso Regression and Neural Network in predicting sales demand for both datasets. Optimisation techniques like PSO further refined predictions, highlighting the importance of selecting appropriate techniques for better performance. This research highlights Ridge Regression's potential for accurate sales forecasting.



Northumbria University
NEWCASTLE