Name :- K.L.V Jayaram
Course Code: SPIC4A07
Reg.No: 19221314
Guide: Dr. A. Moorthy
Dept: CSE
Project No: SSE/26/10/314-1

Title-1:-

Evaluating the Precision of Random forest vs Support Vector Machines for Forecasting Job Rescission in the Industry

Introduction:-

Paragraph 1:-

1) Definition:-

* This study compares Random forest and Support Vector Machine algorithms to determine the most effective approch for predicting the Job Rescission in the Industry.

2) Importance in Today's World:-

In the current dynamic economic environment, industries face uncertainties that can lead to the rescission of job offers. Accurately forecasting such events is vital for organizations to maintain their reputation and for job candidates to manage their career expectations effectively.

Applications of Reasearch:

* Human Resource Managent:
  Enhancing the reliability of hiring processes by anticipating potential rescissions.
* Risk Management:
  Identifying economic or organizational indicators that may lead to offer withdrawals.
* Policy Development:
  Assisting in the creation of guidelines to minimize the impact of rescissions on both companies and prospective employees.

Paragraph 2:-

1) Total Number of arcticles Published in the Past 5 years.

→ Specific data on the number of articles focused solely on forecasting job rescission using Random Forest (RF) and Support Vector Machines (SVM) is limited. However, numerous studies have applied these machine learning techniques to related fields, such as predicting employee turnover and job market trends.

Paragraph-3:-

1) Existing Experience in Research:

→ Studied so many articles regarding on this topic and visited so many websities on the issue & discussed with teachers on this topic. Processing such as data augmentation & normalization are crucial to Ensure that model can learn robust features.

2) Aim of Study:

To improve the accuracy of predicting the Job recission in the industry by analytically Evaluating the performance of RandomForest and SVM.

Materials and Methods:-

Paragraph-1:-

Study setting - SIMATS [SIMATS Engineering]

No. of Groups : 2

Group 1: RF

Group 2: SVM

* this step took advantage of the data exploration and Quality Verification made earlier to create the final data.

2. Most Cited Articles and Findings:

* HVAC Fault Detection: A study introduced a hybrid approach combining RF and SVM for fault detection in HVAC systems, achieving a prediction accuracy high.

* Sentiment Text Analysis: Research comparing SVM, XGBoost and RF for sentiment analysis found that RF outperformed the other models, suggesting its robustness in handling complex datasets.

3) Best Study in Our opinion:

* Recession Prediction Using Multiple Machine Learning Methods and Historical Economic Data - Philip Mackay. -University of Lincoln - Mubeen Ghafoor - University of Lincoln - November 10th, 2023.

Sample size: 20

Total size: 40

Paragraph -2:-

Same Graph:-

Pseudo Code:-

Random Forest Algorithm:

→ Import Libraries
→ Load data
→ Preprocess data
→ Split data
→ Train model
→ Compare model
→ Result.

Paragraph-3:-

Pseudo Code for Support Vector Machine (SVM):

→ Import Libraries
→ Process the data
→ Split the dataset into features
→ Compute models.
→ update weights
→ Compare Nodes
→ Result

Paragraph-4:

Testing Setup:

→ Data preparation
→ Feature Engineering
→ Data Splitting
→ Model Training
→ Class Validation
→ Comparision.

Data Collection:- * Kaggle * IEEE Explorer

Results & Discussion:

* In related studies, both RF and svm have demonstrated high predictive capabilities.
* the choice between them often depends on the specific characteristics of the dataset and the nature of the prediction task.

Limitations:

→ Data Availability
→ Feature selection
→ Model Interpretability

Future Scope:

→ Hybrid Modeling
→ Feature Engineering
→ Cross-Industry Analysis

Conclusion:- Evaluating the precision of Random Forest and Support Vector Machines in fore casting job rescission is essential for improving workforce stability and Planning.

Group Statistics

	GROUP	N	Mean	Std. Deviation	Std. Error Mean
ACCURACY	RF	10	97.0790	.94658	.29933
	SVM	10	94.7200	.76056	.24051

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
ACCURACY	Equal variances assumed	.445	.513	6.143	18	.000	2.35900	.38399	1.55227	3.16573
	Equal variances not assumed			6.143	17.202	.000	2.35900	.38399	1.54958	3.16842