

A photograph of two men in a dimly lit office environment. They are seated at a desk, looking at a laptop screen. The screen displays a financial chart with a line graph and a bar chart. The man in the foreground is wearing glasses and a blue shirt, and is pointing at the screen with a pen. The man in the background is also looking at the screen. The overall atmosphere is professional and focused.

Behavioural Biases and Predictive Modelling of Mutual Fund Decisions

This study explores how behavioural and demographic factors influence mutual fund investment decisions among Indian expatriates in Ireland. Using data from 400 Indian investors, the research applies machine learning classification models to predict investment behaviour while conducting a cost-profit analysis to evaluate financial implications.

The study checks three ideas: if behavioural and demographic factors really affect investment choices, if machine learning models can accurately predict mutual fund investment behaviour, and if using a cost-profit approach makes classification models more useful in behavioural finance.



by **Rakesh K**

Understanding Investor Behaviour in Global Finance



Transnational Investment Environment

Indians living abroad represent a unique demographic whose investment behaviour is shaped by dual socio-economic contexts—their resident country and country of origin.



Behavioural Finance Insights

Psychological and emotional factors often drive investment decisions, leading to deviations from traditional rational models through biases like overconfidence, herding, and loss aversion.

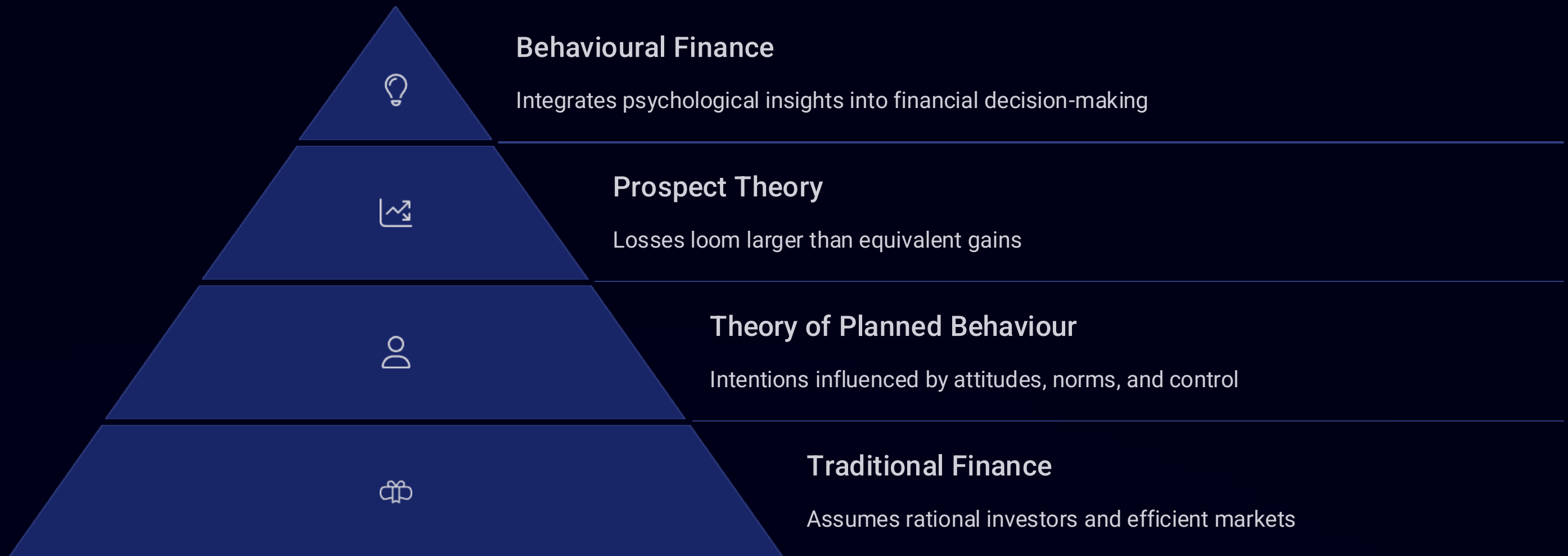


Machine Learning Applications

Researchers have begun to model and predict complex behavioural patterns using machine learning, offering new lenses to interpret investor sentiment and decision biases.



Theoretical Foundations of Behavioural Finance



Traditional financial theories posit that investors are rational agents who make decisions based on all available information. However, empirical observations reveal that investors often deviate from rationality due to psychological biases and heuristics. Prospect Theory suggests individuals evaluate potential gains and losses compared to a reference point, showing loss aversion. The Theory of Planned Behaviour highlights how social influences and perceived control affect investment actions.

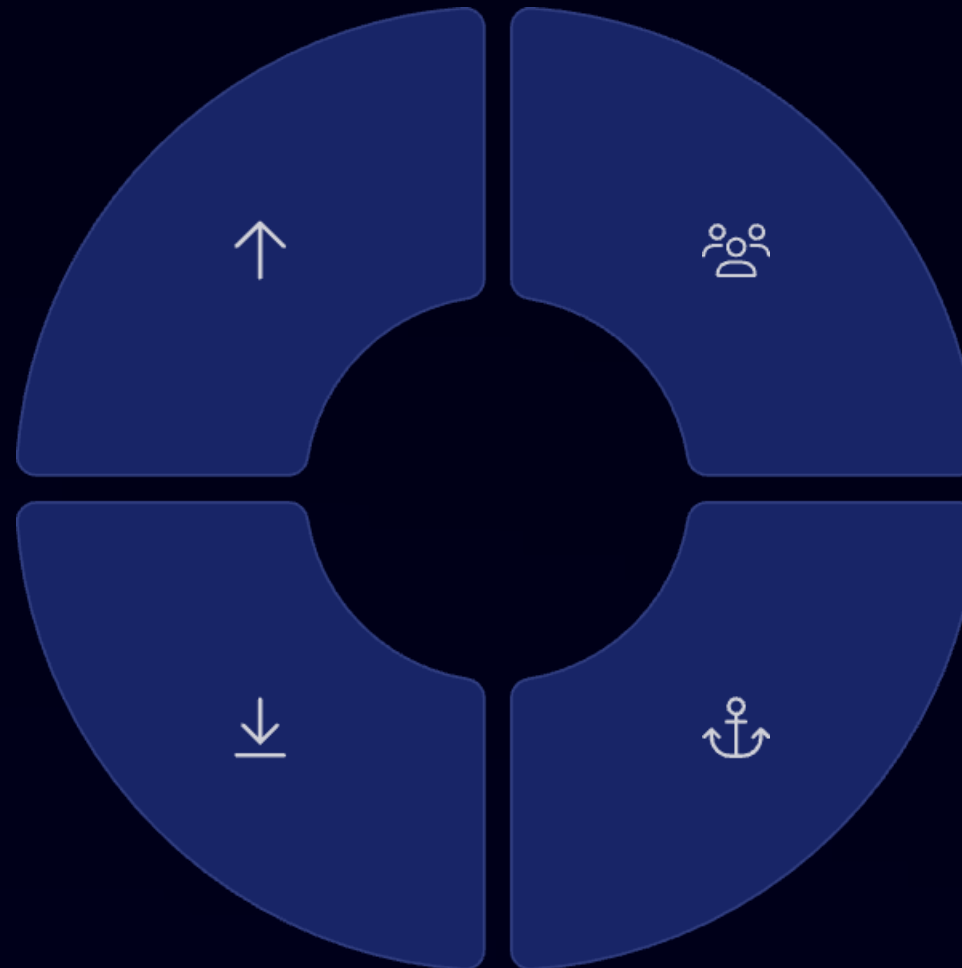
Common Behavioural Biases in Investment Decisions

Overconfidence Bias

Investors overestimate their knowledge and predictive abilities, often resulting in excessive trading and suboptimal returns.

Loss Aversion

The pain of losses is felt more intensely than the pleasure of equivalent gains, leading investors to avoid realizing losses.



Herd Behaviour

Individuals mimic the actions of larger groups, potentially driving asset prices away from intrinsic values and leading to market bubbles.

Anchoring Bias

Investors rely heavily on initial information or specific reference points, potentially ignoring subsequent relevant data.

These biases are not isolated; they often interact, compounding their effects on investment decisions. Understanding these psychological factors is crucial for developing effective investment strategies and advisory services.

Research Methodology and Data Collection



Population and Sampling



Target population included Indian nationals living in Ireland who have made investment decisions pertaining to India, with 500 respondents selected through purposive sampling.

Data Collection



Structured questionnaires captured demographic information, behavioural traits, and investment preferences using Likert scales from 1-5.

Data Preprocessing



Robust preprocessing included handling missing data, encoding categorical variables, feature scaling, and train-test splitting with a 70:30 ratio.

Analysis Techniques



Various supervised machine learning classification algorithms were applied, with performance evaluated using accuracy, precision, recall, and F1 score.

Demographic Profile of Indian Investors in Ireland

Age and Education

Most respondents fall within the 31-45 years age group, followed by those aged ≤ 30 . Graduates and professionals form the highest educational qualifications, suggesting a well-educated investor base.

Location and Family Status

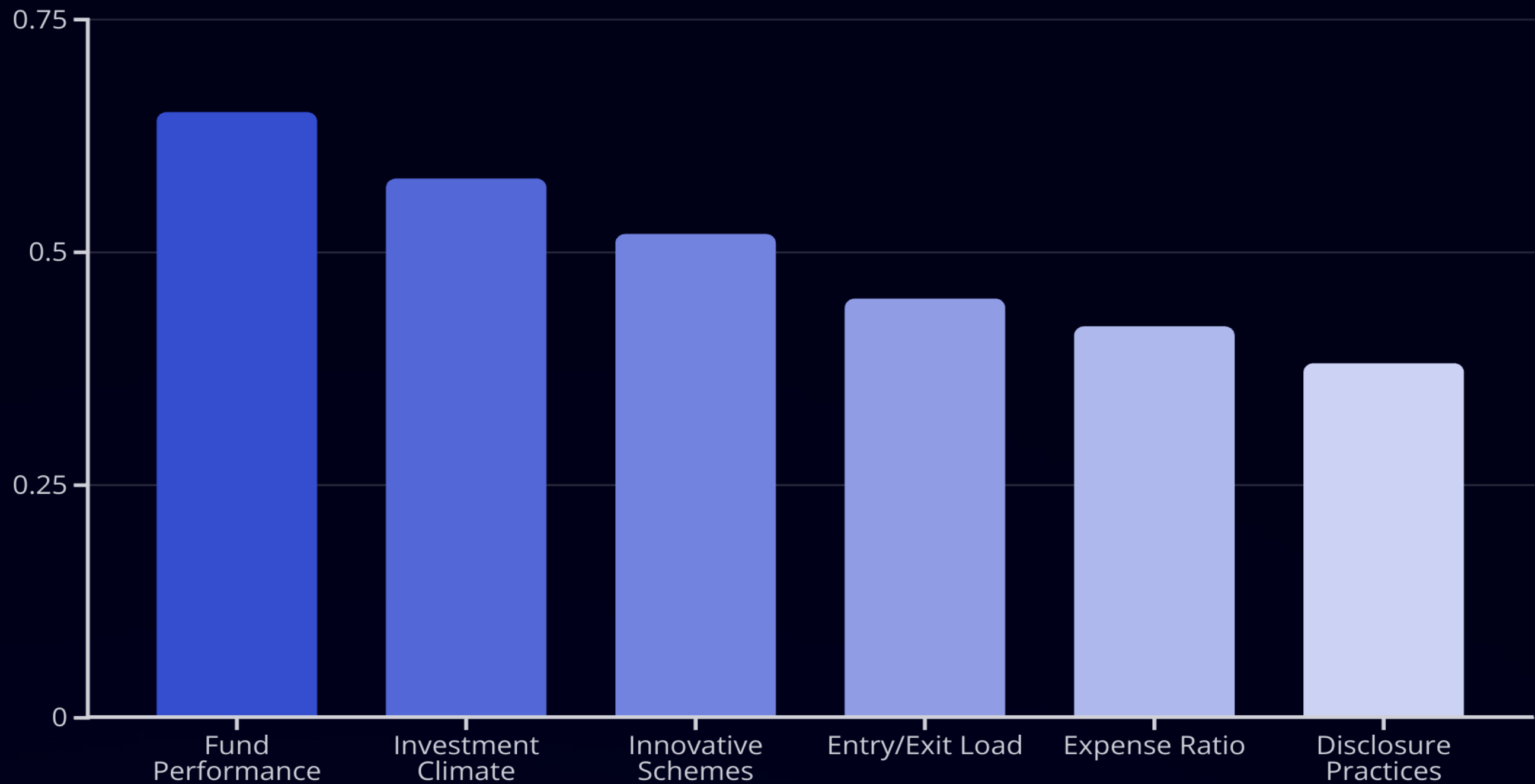
A considerable number are in urban areas, with fewer participants from rural regions. Most respondents are married, and small to mid-sized families (≤ 5 members) dominate the sample.

Occupation and Income

A large segment works in the IT and ITES sector, followed by individuals from other service sectors and public organizations. Most participants earn between €30,000 and €60,000 annually, showing a moderate to upper-middle income group.

These distributions help contextualize the investment behaviour analysed in later sections, reflecting a young, urban, educated, and professionally active demographic that shapes their approach to mutual fund investments.

Factors Influencing Mutual Fund Investment Decisions



Machine Learning Model Performance

Model	Accuracy	True Positives	False Positives	True Negatives	False Negatives
Logistic Regression	0.508	32	33	29	26
Decision Tree	0.575	34	27	35	24
Random Forest	0.583	30	22	40	28
Support Vector Machine	0.55	27	23	39	31

Among the models tested, Random Forest showed the highest predictive accuracy at 58.3%, followed by Decision Tree at 57.5%. While these accuracy levels are modest, they represent an initial exploratory benchmark for modeling complex investor behaviours using behavioural and demographic variables.

The ROC curve yielded an AUC score of 0.65 for the Random Forest model, indicating moderate discriminatory power. The confusion matrix reveals balanced classification between mutual fund investors and non-investors, though with room for improvement.

Cost-Profit Analysis of Prediction Models

€1,420

Decision Tree Net Profit

Highest financial return despite not having the best accuracy

€1,140

Random Forest Net Profit

Second highest return with best accuracy (58.3%)

€600

Logistic Regression Net Profit

Lower return despite similar true positive count

€550

SVM Net Profit

Lowest financial return among all models

While traditional classification metrics are essential for evaluating technical performance, they don't fully capture real-world impact. This cost-profit analysis translates classification outcomes into monetary terms, showing that the Decision Tree model generated the highest net profit despite the Random Forest having better accuracy. This highlights the importance of considering financial implications alongside technical metrics when evaluating models for investment behavior prediction.

Conclusions and Future Research Directions

Key Findings

The study found modest predictive performance across all models, with Random Forest achieving the highest accuracy at 58.3%. The research partially supports the hypothesis that behavioural and demographic variables influence investment decisions, though their standalone predictive power remains moderate.

This study represents an important step in understanding the complex interplay between behavioural finance and machine learning in the context of expatriate investment decisions, offering valuable insights for financial advisors and policymakers targeting diaspora investors.

Contributions

This research introduces a unique focus on Indian diaspora investors, integrates a cost-profit framework with machine learning outputs, and reinforces the role of behavioural factors in shaping investment decisions. It bridges the gap between technical model performance and real-world financial outcomes.

Future Directions

Future research should expand the sample size across multiple geographies and incorporate dynamic variables such as social media sentiment and real-time financial metrics. Enhanced ensemble models, sentiment-enriched features, and deep learning architectures could improve predictive accuracy.