WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Trend Prediction of Punjab Youth Aspiring to Migrate to Canada: A Machine Learning Perspective

Gurpreet Singh[1], Dr. Gurpreet Singh[2], Karuna Devi[3]

[1]Research Scholar M. Tech-CSE, [2]Professor, [3]Assistant Professor
Department of Computer Sc. & Engg., St. Soldier Institute of Engg. & Technology, Near NIT, Jalandhar

Doi : https://doi.org/10.55248/gengpi.5.0924.2412

ABSTRACT:

Migration has long been a significant phenomenon influencing global demographic patterns, economic landscapes, and social structures. Among the various migration trends, the movement of youth from Punjab, India, to Canada has emerged as a particularly noteworthy trend over recent decades. This migration trend is driven by various factors, including educational opportunities, better employment prospects, and the promise of an improved quality of life. The aspiration to migrate to Canada among Punjab's youth has not only socio-economic implications but also cultural and familial impacts, shaping the future of both regions involved. Machine learning can analyze large datasets from various sources such as immigration records, social media activity, and economic reports to identify trends and patterns. This analysis can reveal the underlying factors driving migration.

Keywords: ML, CCI, ICI, IRF

1. Introduction:

In recent years, advancements in data analytics and machine learning have provided powerful tools to analyze and predict migration trends. Machine learning, with its capability to process vast amounts of data and uncover hidden patterns, offers a robust framework for understanding the complex dynamics behind the migration aspirations of Punjab's youth. By leveraging various data sources such as demographic information, socio-economic indicators, and historical migration data, machine learning models can provide valuable insights into future migration trends.

This research aims to explore the application of machine learning techniques to predict the trend of Punjab youth aspiring to migrate to Canada. Through the analysis of relevant data and the development of predictive models, we seek to identify key factors influencing this migration trend and forecast future patterns. This research not only contributes to the academic understanding of migration dynamics but also provides practical implications for policymakers, educators, and stakeholders involved in managing and supporting the migration process. By understanding the predictive factors and future trends, strategies can be developed to address the aspirations and challenges faced by Punjab's youth, ensuring a balanced and informed approach to migration.

2. Machine Learning Perspective

Machine learning can predict the trend of Punjab youth aspiring to migrate to Canada by leveraging a comprehensive process that includes data collection, preprocessing, model training, and evaluation. Initially, data is gathered from various sources such as surveys, historical migration records, social media discussions, and economic indicators. This raw data is then cleaned, standardized, and transformed into useful features for analysis. Through exploratory data analysis, patterns and correlations are identified to inform the model selection. Suitable models, such as regression for predicting continuous outcomes, classification for categorizing migration aspirations, or time series models for forecasting trends, are chosen and trained on the preprocessed data. The training process involves splitting the data into training and testing sets and using cross-validation to fine-tune the model's parameters. The model's performance is assessed using metrics like accuracy, precision, and mean squared error to ensure reliability. Once validated, the model is deployed to make predictions on new data, integrated into applications for real-time analysis. Continuous monitoring and updating of the model with new data ensure sustained accuracy, helping policymakers and stakeholders understand and respond to the migration aspirations of Punjab youth.

3. Machine Learning Algorithms

Machine learning algorithms like Random Forest, Gradient Boosting, and Long Short-Term Memory (LSTM) can significantly enhance the prediction of trends among Punjab youth aspiring to migrate to Canada. Each algorithm offers unique strengths tailored to different aspects of the prediction task.

3.1 Improved Random Forest

Improved Random Forest is an ensemble learning method that constructs multiple decision trees during training and outputs the mode of the classes for classification tasks. It's particularly useful for handling large datasets with many features and can effectively manage missing values and outliers. In the context of predicting migration trends, Improved Random Forest can analyze various factors such as age, education level, family income, and employment status to determine the likelihood of a youth aspiring to migrate. Its ability to handle categorical and numerical data makes it versatile for this application.

3.2 Gradient Boosting

Gradient Boosting builds models sequentially, each new model correcting the errors of its predecessor. This approach results in a powerful predictive model that can capture complex patterns in the data. For predicting migration aspirations, Gradient Boosting can combine different features and capture intricate relationships between them, providing high accuracy. It's particularly effective in scenarios where the data is noisy or has many interdependent variables, making it well-suited for understanding multifaceted socio-economic factors influencing migration decisions.

3.3 Long Short-Term Memory (LSTM)

LSTM is a type of recurrent neural network (RNN) designed to model sequential data. It is particularly adept at capturing long-term dependencies and temporal patterns, making it ideal for time series prediction. In the context of migration trends, LSTM can analyze historical data on migration rates, economic indicators, and policy changes over time to forecast future trends. By considering the temporal dimension, LSTM can provide insights into how past events and trends influence future migration aspirations.

4. Objectives of the study

1. To analyze historical data to identify and understand the key factors driving the migration aspirations of Punjabi youth to Canada.

2. To develop machine learning models to accurately predict future trends in migration aspirations, focusing on both individual-level predictions and broader population trends.

3. To evaluate the influence of changes in immigration policies, economic conditions, and educational opportunities on migration aspirations, providing insights for policymakers.

5. Results and Discussion

Dataset Used

ID	AGE	GENDER	EDUCATIONAL LEVEL	OCCUPATION	INCOME (MONTHLY)	SOCIAL MEDIA INFUENCE (1-10)	FRIENDS/ RELATIVES IN CANADA	IELTS SCORE	MIGRATION INTENSION (1-10)	YEAR	MIGRATED TO CANADA
1	18	MALE	+2	STUDENT	15000	8	YES	6.5	7	2022	YES
2	20	MALE	GRADUATION	STUDENT	25000	6	NO	7	7	0	NO
3	18	FEMALE	+2	STUDENT	16000	7	NO	6	8	2023	YES
4	21	FEMALE	GRADUATION	STUDENT	17000	9	YES	6.5	6	0	NO
5	27	MALE	GRADUATION	TECHNICIAN	26000	8	YES	7.5	9	2020	YES
6	26	FEMALE	MASTERS	TEACHER	18000	7	NO	7	7	2021	YES
7	20	FEMALE	GRADUATION	STUDENT	20000	8	NO	6	8	0	NO
8	28	MALE	GRADUATION	ENGINEER	25000	9	NO	7	8	2022	YES
9	21	MALE	GRADUATION	STUDENT	19000	8	YES	6.5	9	2022	YES
10	29	FEMALE	GRADUATION	NURSING	16000	5	NO	7	7	0	NO
11	25	MALE	GRADUATION	LAWER	20000	6	NO	7.5	8	2021	YES

12	18	FEMALE	+2	STUDENT	15000	7	YES	6	7	2023	YES
13	24	FEMALE	GRADUATION	TEACHER	14000	8	YES	7	7	0	NO
14	22	MALE	GRADUATION	STUDENT	16000	7	NO	6.5	8	2023	YES
15	25	MALE	MASTERS	ENGINEER	26000	5	YES	7	7	2022	YES
16	21	FEMALE	GRADUATION	STUDENT	17000	9	YES	6.5	6	0	NO
17	27	MALE	GRADUATION	TECHNICIAN	26000	8	YES	7.5	9	2020	YES
18	26	FEMALE	MASTERS	TEACHER	18000	7	NO	7	7	2021	YES
19	20	FEMALE	GRADUATION	STUDENT	20000	8	NO	6	8	0	NO
20	28	MALE	GRADUATION	ENGINEER	25000	9	NO	7	8	2022	YES
21	21	MALE	GRADUATION	STUDENT	19000	8	YES	6.5	9	2022	YES

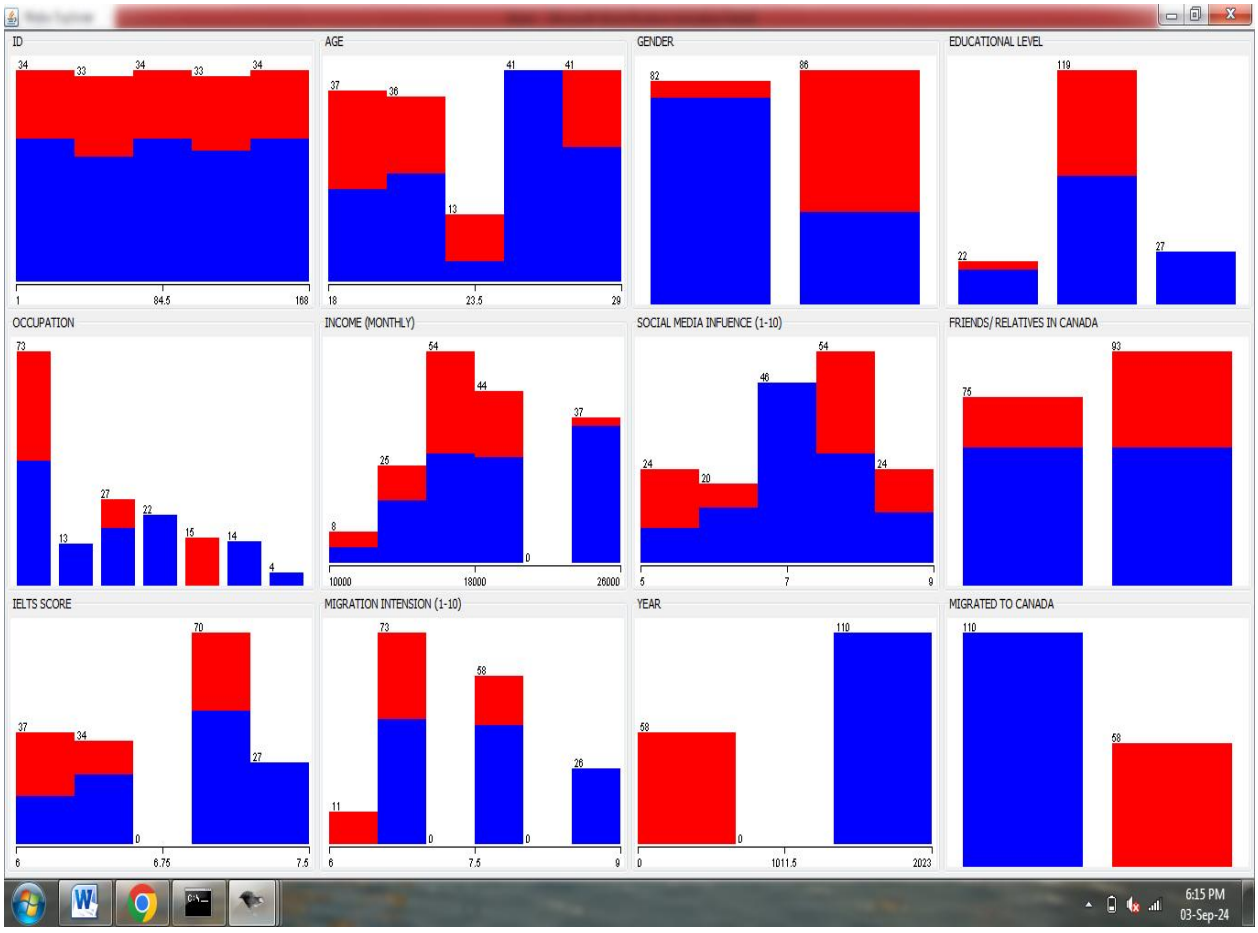

Fig. 1: Visual representation of Punjab Youth Migration Dataset

Table 1: Comparison of Improved Random Forest, Gradient Boosting, and LSTM Algorithms Based on Correctly and Incorrectly Classified Instances

Parameters	**Gradient Boosting Algorithm**	**Improved Random Forest Algorithm**	**LSTM Algorithm**
Correctly Classified Instances	70.23	95.23	68.31
Incorrectly Classified Instances	29.76	4.76	31.68

Comparison of Improved Random Forest Algorithm, Gradient Boosting Algorithm and LSTM Algorithm Based on Correctly and Incorrectly Classified Instances

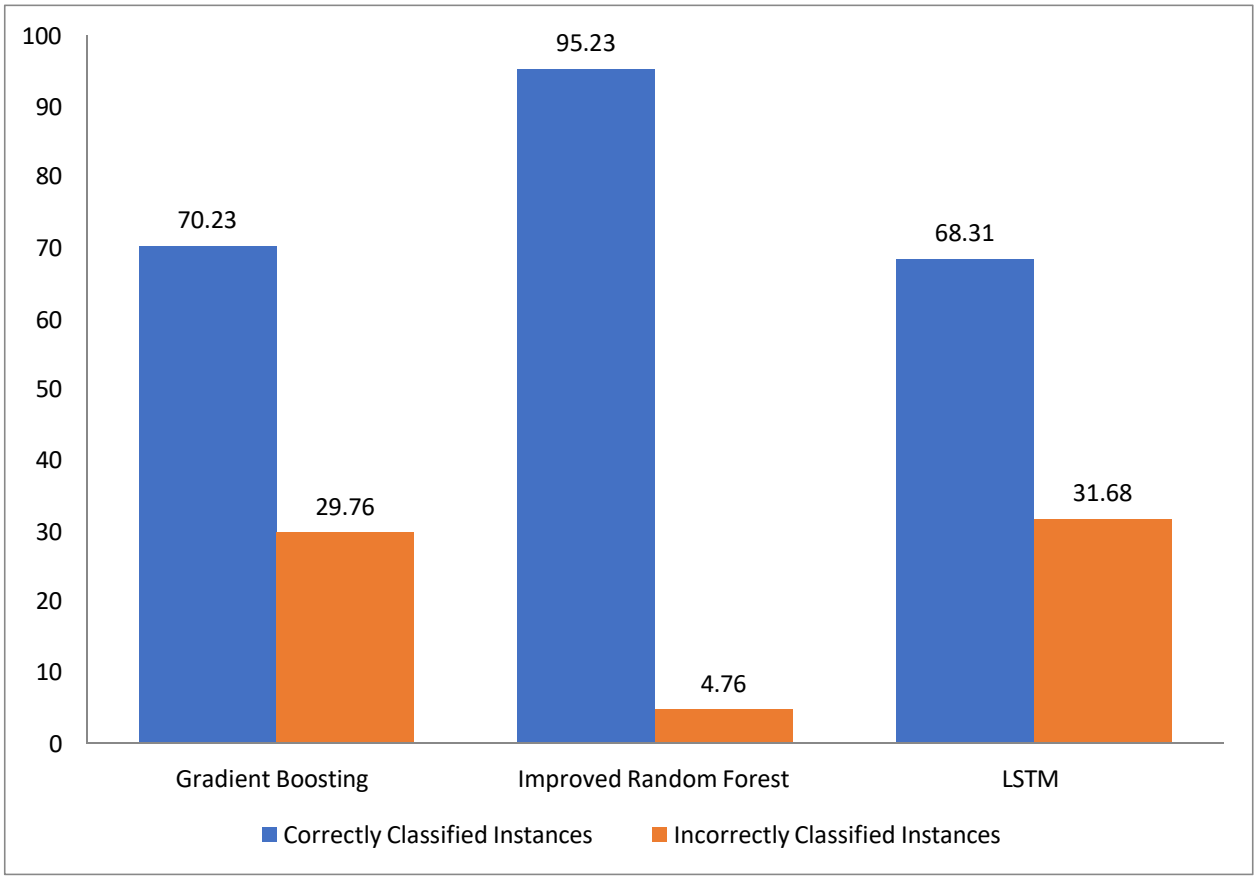

Fig. 2: Comparison of Improved Random Forest Algorithm, Gradient Boosting Algorithm and LSTM Algorithm Based on Correctly and Incorrectly Classified Instances

6. Conclusion

Based on the provided figures, the study focused on predicting the migration trend of Punjab youth aspiring to move to Canada using three machine learning algorithms: an Improved Random Forest Algorithm, a Gradient Boosting Algorithm, and an LSTM Algorithm. The accuracy of these algorithms was measured by the percentage of correctly and incorrectly classified instances.

The Improved Random Forest Algorithm demonstrated the highest classification accuracy, with 95.23% of instances correctly classified and only 4.76% incorrectly classified. This indicates that the Improved Random Forest is highly effective in predicting migration trends among Punjab youth, making it the most reliable model among the three.

In contrast, the Gradient Boosting Algorithm and the LSTM Algorithm performed less effectively. The Gradient Boosting Algorithm correctly classified 70.23% of instances, with a significant 29.76% of instances incorrectly classified. Similarly, the LSTM Algorithm showed a lower accuracy, with 68.31% of instances correctly classified and 31.68% incorrectly classified.

These results suggest that while the Improved Random Forest Algorithm provides a strong predictive performance, the Gradient Boosting and LSTM Algorithms require further optimization to enhance their accuracy. The high performance of the Improved Random Forest Algorithm makes it a preferable choice for applications requiring high precision in trend prediction, particularly in the context of migration studies. The study's findings underscore the importance of selecting the appropriate machine learning model based on the specific requirements of accuracy and reliability.

References

[1] Ruchir Agarwal, Ina Ganguli , Patrick Gaul,Geoff Smith, " Why U.S. immigration matters for the global advancement of science",ELSEVIER, ScienceDirect, 2023,

[2] Yasmin Y. Ortiga ,Michael Joseph, Romeo Luis Macabasag, "Clocking out: Nurses refusing to work in a time of pandemic", ELSEVIER, ScienceDirect, 2022.

[3] Amanda Miller, "Development through vocational education. The lived experiences of young people at a vocational education, training restaurant in Siem Reap, Cambodia", Heliyon, ScienceDirect, 2020.

[4] Reem AlKharouf, Ali Shehadeh , Areej Alrefaee , Odey Alshboul , " Integrative strategies for social inclusion and equity: Enhancing refugee access to higher education in Jordan", Heliyon, ScienceDirect, 2024.

[5] Regina Gairal Casado, David Golightly, Karen Laing , Roberto Palacin , Liz Todd, "Children, Young people and Mobility as a Service: Opportunities and barriers for future mobility", Transportation Research Interdisciplinary Perspectives, ELSEVIER, ScienceDirect, 2020.

[6] Leysen, Bert Verhaeghe, Pieter-Paul, "Searching for migration: estimating Japanese migration to Europe with Google Trends data", Quality & Quantity, 2023.

[7] Yixi Lu, "Studying Abroad and Migration Motivations: A Case Study of Chinese Students At The University Of Saskachewan, Canada", Thesis, College of Graduate Studies and Research, provided by University of Saskatchewan's Research Archive, 2006.

[8] Marina Dabic, Bozidar Vlacic, Justin Paul, LeoPaul Dana, Sreevas Sahasranamam, Beata Glinka "Immigrant Entrepreneurship: A Review and Research Agenda", Journal of Business Research, ELSEVIER, ScienceDirece, 2020.

[9] Sri Purwatiningsih, "International Migration and Aspirations among Young People: Who Desires to Migrate?", Populasi Volume 29 Issue 2 2021.

[10] Samuel Hall," Youth, Migration and Development: A New Lens for Critical Times", KNOMAD, Paper No 41, World Bank, Washington, DC. 2021.

[11] Santosh Kumar Mahato, Devi Prasad Paudel, "Motivating Factors of Youth Selecting Labor Migration Rather Than Higher Education: A Systematic Review", **Interdisciplinary Research in Education Volume 7, Issue 1, 2022.**

[12] Anneke Newman, "The influence of migration on the educational aspirations of young men in northern Senegal: Implications for policy", International Journal of Educational Development, , ELSEVIER,ScienceDirece, 2018.

[13] Carol Graham, Julia R. Pozuelo, "Do high aspirations lead to better outcomes? Evidence from a longitudinal survey of adolescents in Peru", Journal of Population Economics (2023).

[14] Puja Chatrath, "Challenges Faced By New Migrants Entering Existing Workplace Communities in the UK", University of Stirling Management school, 2022.