

Predicting Employee Promotion using Machine Learning

K. Durves Mohideen¹ and S. Prasanna²

MCA Student, Department of Computer Applications¹

Professor, Department of Computer Applications²

durvesmohideen68@gmail.com and prasanna.scs@velsuniv.ac.in

Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India

Abstract: *Training and development are key components of professional development for people to improve their capacity. Professional development programs are typically organized around personal information like as background, personal goals, and work experience, as well as corporate objectives and job requirements. Individual employee classification is required to promote tailored training in the professional development process. As a result, this study provides a classification approach for employee classification in order to facilitate tailored training in enterprises. Machine learning methods such as Decision Tree, Random Forest, and Support Vector Machine are investigated. To cope with imbalance data, the Synthetic Minority Oversampling Technique (SMOTE) approach is applied. In this work, the open data form kaggle is used. The training and testing data are combined to generate the data for technique validation. There are three gropes: 80:20, 70:30, and 60:40. According to the classification results, the SMOTE can increase classification performance for all classifiers. Furthermore, random forest has the highest categorization accuracy.*

Keywords: Professional development

I. INTRODUCTION

A technique for managing human resources in an organization is called human resource management. Training and development are a key element of human resource management. Training is a quick process the attaches the fundamental knowledge and abilities needed for any job. In contrast, development is a lengthy process that concentrates on individual or personal growth to enhance performance for the demands of future jobs, including conceptual or theoretical inputs, perspective strategic thinking, and behavioral aspects, such as leadership skills, team management, etc. Employees must be able to tackle difficult or complex tasks during the development process. So, it invests time in a lengthy process. A process of growth is called professional development (PD). Through learning and applying procedures, it makes use of prior information, abilities, and experience to enhance personal capacity. There are numerous methods to learn and grow. Specialized knowledge, abilities, techniques, and skills. Since job recruitment is no longer necessary to meet company objectives or job position requirements, professional development is now a key component of many organizations. Many things can occur during the training and development process, such as personnel or organizational problems. Training and development processes, particularly for organizational problems, are typically expensive, time-consuming, and fraught with difficulties in selecting competent people. An organization will often select personnel with a high professional growth process based on these evidences. As a result, individuals' skills and prior knowledge must be evaluated and compared to job criteria. However, every employee needs a unique training and development procedure. The purpose of this paper is to provide a suitable approach for selecting the right individual for the correct training and development process in terms of personal goals and corporate objectives.

II. LITERATURE SURVEY

[1] Y. Asim, B. Raza, A. K. Malik, S. Rathore and A. Bilal, "Improving the Performance of Professional Blogger's Classification", 2018. International Conference on Computing, Mathematics and Engineering Technologies (I COMET), Sukkur, Blogging is a good method for writing internet content, and those who participate in this activity are known as bloggers. A blogger can be classed into classes based on characteristics such as educational background, cultural background, and topical interests. Bloggers may choose this job due to a variety of circumstances (influencing characteristics). The topic of this study is the classification of professional bloggers and the identification of such relevant criteria. For the binary classification problem of a blogger's dataset, we employed an Artificial Neural Network. For factor identification, the Predictive Apriori association rule mining technique is utilized. In this research, the outcomes of an Artificial Neural Network are compared to those of the Random Forest and Nearest-Neighbor algorithms. It has been discovered that Artificial With 87% accuracy and 86.9% F-measure, Neural Network surpasses Random Forest (RF) and Nearest Neighbour (IB1) algorithms. The factor identification findings are compared to the results of the Alternate Decision Tree (ADTree) technique. Both the AD Tree and the Predictive Apriori algorithms yielded the same results for predictive performance measurements...

[2] T.W Ramdhani, B. Purwandari and Y. Ruldeviyani, "The Use of Data Mining Classification Technique to Fill in Structural Positions in Bogor Local Government" 2016 International conference on Advanced computer Science and Information Systems (ICACSIS), Malang 2016 Bogor's human resources are administered by the human resources and training division, known as Badan Kepegawaian Pendidikan dan Pelatihan (BKPP). BKPP forms the Badan Pertimbangan Jabatan dan

Kepangkatan (Baperjakat) team, which is in charge of promoting, rotating, and dismissing local government personnel from structural posts below Echelon IIA. Baperjakat is having difficulty creating a draft of structural government roles. Despite the fact that BKPP has a human resources information system called SIMPEG, these operations were carried out manually. The primary goal of this study is to find patterns for filling structural positions in the Bogor Local Government. To detect filling structural position patterns, 62 classification methods were tested utilizing three data mining tools, seven data sets, and seven human resource qualities. The best algorithm in the echelon class according to the classification procedure is Classification Rule with Unbiased Interaction Selection and Estimation (CRUISE). For every echelon level, it has an average accuracy of 95.7%...

III. METHODOLOGY SECTION

Methodology refers to the systematic approach or set of principles and practices used by researchers, scientists, or professionals to conduct their work or investigations. It outlines the procedures, techniques, and tools employed to gather, analyze, and interpret data or information in a particular field of study or endeavor. Methodology plays a crucial role in ensuring the rigor, reliability, and validity of research or project outcomes.

In various disciplines, methodologies can vary significantly based on the nature of the research questions, the type of data being collected, and the objectives of the study. For instance, in scientific research, methodologies often involve experiments, observations, and data analysis using specific techniques and instruments. In social sciences, methodologies may include surveys, interviews, case studies, or content analysis.

Key components of a methodology typically include:

- **Research Design:** This outlines the overall plan or strategy for conducting the study, including the selection of participants, data collection methods, and analytical techniques.
- **Data Collection Methods:** These are the specific techniques used to gather information or data relevant to the research objectives. Common methods include surveys, interviews, experiments, observations, and document analysis.
- **Ethical Considerations:** Methodologies often include provisions for ensuring that research is conducted ethically and with respect for the rights and welfare participants. This may involve obtaining informed consent, maintaining confidentiality, and addressing potential risks or conflicts of interest.

- **Validity and Reliability:** Methodologies typically address issues of validity (whether the research accurately measures what it intends to measure) and reliability (the consistency and stability of research findings over time and across different contexts).

Data collection.

The data collection module in the Employee Promotion System is a crucial component responsible for gathering, storing, and managing employee-related data that will be used for training and deploying machine learning models. This module serves as the foundation for making data-driven promotion decisions using decision tree and Naive Bayes algorithms within the Flask framework. It provides structured and unstructured data about employees, such as their performance reviews, skills, experience, and historical promotion data.

Data pre-processing.

Once data is extracted, it undergoes pre-processing to clean and transform it into a usable format. This includes handling missing values, normalizing data, and encoding categorical variables. Data pre-processing ensures that the machine learning models receive high-quality input.

IV. SYSTEM ARCHITECTURE

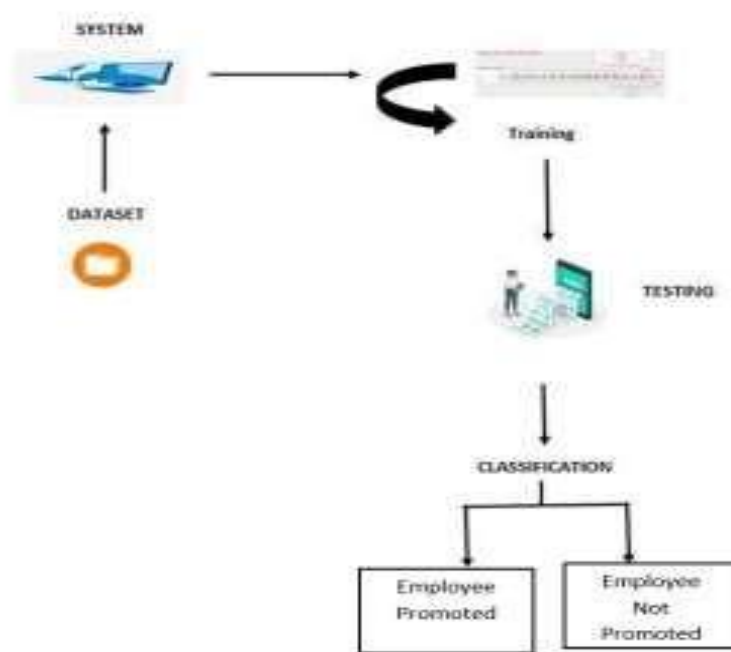


Fig 1: Frame work of proposed method.

Model implementation.

This module serves as the core decision-making engine, leveraging these models to predict and facilitate employee promotions. The module initiates the training process for the machine learning models, using historical employee data as input. The decision tree algorithm and Naive Bayes classifier are trained separately to capture different aspects of promotion prediction. Once trained and validated, the models are serialized to a format that allows for easy storage and retrieval. This enables the models to be loaded quickly when making predictions.

The serialized model objects, in pickle format, are then stored as files in a designated location within the system's file system.

The Flask application integrates with the previously saved machine learning models (decision tree and Naive Bayes) by loading them into memory. This allows the models to be used for real-time prediction when user input is received.

Prediction

The prediction module is an integral part of an employee promotion system designed to assist organizations in making data-driven decisions about promoting employees. This module leverages machine learning algorithms, specifically Decision Trees and Naive Bayes, to predict whether an employee should be promoted based on various input features. The Decision Tree and Naive Bayes models are then applied to the pre-processed data to make predictions about whether an employee should be promoted. The predictions are returned as binary outcomes (e.g., "Promote" or "Not Promote") or as probability scores, depending on the specific requirements.

V. CONCLUSION

In this study, we created a web application that determines if an individual is eligible for promotion or not. We determined this based on a few parameters, including the employee's experience, previous year ratings, KPIs performed, average training score, and number of trainings attended. This web program also analyzed basic characteristics such as an employee's gender, age, region, education, and department to compute and reveal whether or not the individual is suitable for promotion. We employed machine learning models such as Decision Tree, Random Forest, and Support Vector Machine (SVM) to do this.

REFERENCES

- [1]. Y. Asim, B. Raza, A. K. Malik, S. Rathore and A. Bilal, "Improving the Performance of Professional Blogger's Classification", 2018. International Conference on Computing, Mathematics and Engineering Technologies (I COMET), Sukkur, 2018, pp.1-6.
- [2]. T.W Ramdhani, B. Purwandari and Y. Ruldeviyani, "The Use of Data Mining Classification Technique to Fill in Structural Positions in Bogor Local Government" 2016 International conference on Advanced computer Science and Information Systems (ICACSIS), Malang 2016.
- [3]. Sisodia, D. S., Vishwakarma, S., & Pujahari, A. (2017, November). Evaluation of machine learning models for employee churn prediction. In 2017 international conference on inventive computing and informatics (icici) (pp. 1016-1020). IEEE.
- [4]. Long, Y., Liu, J., Fang, M., Wang, T., & Jiang, W. (2018, May). Prediction of employee promotion based on personal basic features and post features. In Proceedings of the International Conference on Data Processing and Applications (pp. 5-10).
- [5]. Şahinbaş, K. (2022, April). Employee promotion prediction by using machine learning algorithms for imbalanced dataset. In 2022 2nd International Conference on Computing and Machine Intelligence (ICMI) (pp. 1-5). IEEE.
- [6]. Liu, J., Wang, T., Li, J., Huang, J., Yao, F., & He, R. (2019, October). A data-driven analysis of employee promotion: the role of the position of organization. In 2019 IEEE international conference on systems, man and cybernetics (SMC) (pp. 4056-4062). IEEE.
- [7]. M. Husnain, B. Rehman, F. Syed and M. khtar, "Personal and in-store factors influencing impulse buying behavior among generation Y consumers of small cities," Business Perspectives and Research, vol. 7, no. 1, pp. 92-107, 2019
- [8]. R. Akerkar, "Introduction to artificial intelligence," in Artificial Intelligence for Business, Springer, Cham, 2019, pp. 1-18.
- [9]. H. H. Lai, "Study on the influence of employee promotion system on organizational performance," International Journal of Organizational Innovation (Online), vol. 5, no. 1, p. 231, 2012.
- [10]. S. Feng, R. Suri, M. C. H. Chao and U. Koc, "Presenting comparative price promotions vertically or horizontally: Does it matter?," Journal of Business Research, vol. 76, pp. 209-218, 2017