International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 557-560 www.ijaem.net ISSN: 2395-5252

A Machine Learning Approach for Efficient Team Building in the Developer Community

Apoorv Verma, Manas Pandey, Kartik Chaudhary, Aman Dubey , Ramveer Singh

Date of Submission: 05-05-2023 Date of Acceptance: 15-05-2023

ABSTRACT
The Dev-Connect platform serves as a medium for developers worldwide to form teams with like-minded individuals. This platform enables project leaders to create public team member requirements, including essential information such as project descriptions, skills, working hours, and behavioral preferences, to ensure optimal matching. Utilizing machine learning techniques, the platform provides recommendations to project leaders based on similar preferences, facilitating effective team formation.

I. INTRODUCTION

Finding the right team member with compatible understanding, coding behavior, and the necessary skills for a project can be challenging for project creators. To address this difficulty, the Developers Community provides a solution through the Dev-Connect platform. This platform serves as a medium for developers worldwide to form teams with like-minded individuals, making it easier to find suitable team members and fostering a sense of community among developers who share similar mindsets.

Machine Learning:

The Developers Community incorporates machine learning technology to enhance the team formation process. Currently, searching for a reliable project partner and a suitable project based on preferences can be a tedious and time-consuming task. However, by leveraging machine learning, this process becomes easier, faster, and more efficient. The platform utilizes machine learning algorithms to recommend potential team members and projects based on the user's preferences and past selections. This not only facilitates the process of finding the best and most suitable options but also enables knowledge sharing and learning as developers interact with each other. By leveraging machine learning, Dev-Connect empowers developers to explore various options and connects them with relevant opportunities within the community.

II. METHODOLOGY

The methodology employed by the Dev-Ops Developers Community for finding perfect candidates and predicting future trends involves a systematic approach to ensure the effectiveness of the system in identifying and attracting suitable team members.

2.1. Agile Methodology: The Dev-Ops Developers Community adopts the agile methodology for its project management and development process. Agile is a flexible and iterative approach that allows for easy adaptation to changes and prioritizes customer satisfaction. In this context, the system is being developed for the users, and the agile methodology ensures that their needs and feedback are taken into account throughout the development process. It emphasizes collaboration, continuous improvement, and delivering the project in small, incremental cycles known as "sprints."

2.2. Supervised Learning Model: The core of the Dev-Ops Developers Community's system is a machine learning model based on supervised learning. This model is a classification model that predicts the likelihood of a job applicant being hired based on various features and qualifications. These features may include educational background, work experience, skills, certifications, and other relevant factors. The model is trained using historical data that contains information about successful and unsuccessful job applicants.

2.3. Machine Learning Algorithms: To train the supervised learning model, the Dev-Ops Developers Community utilizes machine learning algorithms such as logistic regression, decision trees, or neural networks. The choice of algorithm depends on the specific characteristics of the data

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 557-560 www.ijaem.net ISSN: 2395-5252

and the problem being addressed. These algorithms analyze the input features and patterns within the data to make predictions about the probability of an applicant being hired.

2.4. Feature Engineering and Data Preparation: Prior to training the machine learning model, feature engineering and data preparation steps are undertaken. This involves selecting the relevant features that have a significant impact on the hiring decision and transforming the data into a suitable format for the machine learning algorithms. It may include processes such as data cleaning, normalization, encoding categorical variables, and handling missing values to ensure the quality and consistency of the input data.

III. MACHINE LEARNING MODEL:

The machine learning model used by the Dev-Ops Developers Community is a supervised learning model, specifically a classification model. The model takes as input various features of a job applicant, such as their education, work experience, and skills, and predicts the likelihood of the applicant being hired.

The model is trained on historical data using machine learning algorithms such as logistic regression, decision trees, or neural networks. The specific algorithm used depends on the data and the problem being addressed.

The mathematical approach used by the Developers Community model involves the use of supervised learning algorithms to classify job applicants into two or more categories based on their qualifications and experience. In this approach, the model learns from historical job application data to predict the probability of a candidate being hired for a particular job.

The mathematical approach involves the following steps:

3.1. Feature Extraction: The first step is to extract relevant features from the job application data. This involves identifying the most important variables that can affect the hiring decision, such as education, work experience, and skills.

3.2. Data Preprocessing: The next step is to preprocess the data to remove missing values, handle outliers, and normalize the data. This ensures that the data is consistent and free from any biases that can affect the model's accuracy.

3.3. Model Training: The third step is to train the model on the preprocessed data. This involves feeding the data into a machine learning algorithm that learns from the patterns in the data to predict the outcome of new job applications. Common algorithms used for classification include logistic regression, decision trees, and neural networks.

3.4. Model Evaluation: The final step is to evaluate the performance of the model on a holdout dataset. This involves comparing the predicted outcomes to the actual outcomes to determine the accuracy of the model. Common evaluation metrics include precision, recall, and F1 score.

The mathematical approach used by the model is based on statistical learning theory, which involves using mathematical models to learn from data and make predictions about future events. This approach is widely used in machine learning and has been shown to be effective in a variety of applications, including natural language processing, image recognition, and predictive analytics.

3.5. Goals:

1. Ameliorate deployment frequence ·
2. Achieve faster time to apply
3. Lower failure rate of new releases ·
4. dock supereminent time between fixes
5. Ameliorate mean time to recovery

IV. K-MEANS CLUSTERING:

K- means clustering is a machine learning algorithm used for clustering data points intogroups. It's an unsupervised literacy algorithm that partitions a given dataset into K clusters, where K is the number of clusters specified by the user.The algorithm works by iteratively assigning each data point to the nearest cluster centroid, also streamlining the centroids grounded on the mean of the data points assigned to each cluster. This process continues until the clusters no longer change significantly, or a maximum number of duplications is reached.

4.1. K-MEANS ALGORITHM:

Input: // Set of n items to cluster

D= {d1, d2, d3,……., dn}

// Number of Clusters (Temporary Clusters) Randomly Chosen: k

// In the following representation, K represents a subset of D as a temporary cluster, and C represents the set of centroids for those clusters.

K= {k1, k2, k3,……., kk },

C= {c1, c2, c3,……., ck}

Where k1= {d1}, k2= {d2}, k3= {d3}…… kk= {dk}

And c1=d1, c2=d2, c3=d3,……. ck=dk, here k<=n

Output: // // K is a subset of D as the final cluster and C is the set of centroids of these clusters.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 5 May 2023, pp: 557-560 www.ijaem.net ISSN: 2395-5252

K= {k1, k2, k3,……, kk },
C= {c1, c2, c3,……, ck}
Algorithm:

K-means (D, K, C)
1. Arbitrarily choose k objects from D as the initial cluster centers.
2. Repeat
3. (re) assign each object to the cluster to which the object is the most similar, based on the mean value of the objects in the cluster.
4. Update the cluster means, i.e., calculate the mean value of the objects for each cluster.
5. Until no change.

4.2. LIMITATIONS OF K-MEANS CLUSTERING ALGORITHM:

1. In partitioning based K-means clustering algorithms, the number of clusters (k) needs to be determined beforehand.
2. The algorithm is sensitive to an initial seed selection (starting cluster centroids). Due to selection of initial centroid points it is susceptible to a local optimum and may miss the global optimum
3. It can model only a spherical shape of clusters. Thus the non convex shape of clusters cannot be modeled in center based clustering

4.3. TIME COMPLEXITY OF K-MEANS CLUSTERING ALGORITHM:

To calculate the running time of the K-means algorithm it is necessary to know the number of times each statement runs and the cost of running. But sometimes the number of steps is not known so it has been assumed. For example, the number of times the first statement runs with cost m1 is **q (>=1).** For each q next statement , for **i=1,2,……n** where n is the number of data objects, runs n+1 times with cost m2. For each q and for each n, the next statement runs k+1 times, where k is the number of clusters with cost m3. The 4th statement runs one time for each q and for each n with cost m4. Calculating new mean for each cluster requires k+1 runs for each q with cost m5.
Running time for algorithm is the sum of running time for each statement executed i.e.

T(n) =m1*q+m2*1∑ q (n+1)+m3*1∑ q 1∑ n (k+1)+m4*1∑ q 1∑ n 1+m5*1∑ q (k+1).
=
m1*q+m2*q*(n+1)+m3*q*n*(k+1)+m4*q*n*1+ m5 *q*(k+1).
= m1*q+m2*q*n+ m2*q+m3*q*n*k+ m3*q*n+m4*q*n+m5*q*k+ m5*q
=(m1+m2+m5)*q+(m2+m3+m4)*q*n+m3*q*n*k.

For worst case it will be O(ni) where 2<=I
For best case it will be O(n)
For the average case it will be O(n2).

CONCLUSION:

No doubt, this methodology will boom your production if applied correctly, it is certainly a no match way of carrying out application development. It improves the business performance of application and allows the end user to directly contribute to the application development process. It eliminates all most all conflicts that the development and operation staff used to have while developing the project. Dev-Ops:Developers Community is all about providing a path for faster time to market of new software features and achieving better stability. Right now, there is no chance it is going to be replace rather we will witness it to be a top option for developing dynamic applications that constantly evolve to meet new challenges. Here I have briefly summarized what is Dev-Ops:Developers Community methodology and how and why it has overthrown the previous methodologies. Community for developer where they can come up with new and fresh ideas which can actually solve some real issues. It would not only make the process easier, faster and more efficient but also provide more relevant opportunities and options. The system would also provide information about the more rising and lovable technologies.

REFERENCES:

[1]. Stiegler, Andreas & Loitsch, Claudia & Kaklanis, Nikolaos & Strobbe, Christophe & Votis, Konstantinos & Makris, Lambros & Pavlopoulos, Vissarion & Stavropoulos, Georgios & Kilintzi, Vasilis & Tzovaras, Dimitrios. (2015)Matchmaking Algorithms and Their Evaluation.
[2]. Bleier, Alexander & Eisenbeiss, Maik. (2015) Personalized Online Advertising Effectiveness
[3]. Li, Anita & Mahoney, Amanda & Poling, Alan. (2018). Basic Research in Behavior Analysis. Behavior Analysis: Research and Practice. 18. 117-118. 10.1037/bar0000134.
[4]. Muhammad Usman Tariq, Muhammad Babar, Marc Poulin, Akmal Saeed Khattak, Mohammad Dahman Alshehri and Sarah Kaleem (2021) Human Behavior Analysis Using Intelligent Big Data Analytics.

[5]. Petra Audy Martínek (2021).Mapping methods of research on consumer engagement with brands on social media: A literature review Methodological Innovations January-April 2021: 1–23
[6]. Digital matchmaker: almost half would trust AI to pair them with compatible.Source:https://www.kaspersky.com/about/pressreleases/2021_digital-matchmaker-almost-half-would-trust-ai-to-pair-them-with-someone-compatible
[7]. AI solutions will see greater success by reducing friction and helping solve defined business problems. Source: https://www.tableau.com/reports/data-trends
[8]. How to come up with mobile app Ideas. Varous Mobile App Ideas. Source:https://www.designveloper.com/en/blog/mobile-app-ideas/
[9]. Azure OpenAI Service. Source: https://news.microsoft.com/ignite-2022-book-of-news/