

A Machine Learning Model: Student Verification and Job Placement with Gradient Boosting Machines

E.Sujatha¹, R.G.Sakthivelan², S.Sree Subha³ and Gopirajan PV⁴

¹Department of Computer Science and Engineering, Saveetha Engineering College (Autonomous), Thandalam, Chennai, India,

Email: sujathae@saveetha.ac.in, <https://orcid.org/0000-0003-2500-2937>

²Department of Computer Science and Engineering, Gopalan College of Engineering and Management, Bengaluru, India.

Email: sakthi21245@gmail.com, <https://orcid.org/0000-0002-3444-9821>

³Department of Artificial Intelligence and Machine Learning, Rajalakshmi Engineering College (Autonomous), Chennai,

Email: sreesubha.s@rajalakshmi.edu.in, <https://orcid.org/0000-0002-5937-597x>.

⁴Department of Computational Intelligence, School of Computing, SRM Institute of Science and Technology, Kattankulathur Campus, Chennai, India, 603203

Email: gopivrajan@gmail.com

Abstract— Every year, millions of students graduate, making it difficult to track and process their data. The One Point Student Verification Machine learning Model (OSVMLM) collects rich data on students, including biometric information, secondary education details, list of degrees completed, and detailed information on their last or current degree. In addition, the student's fingerprint is also collected and stored. The government's Aadhar database is used to verify the authenticity of the student. To ease the recruiter, the details collected by OSVMLM were assigned with ranks, information will be sent to the candidates autonomously based on the current job description. This makes it easy for students to identify and apply for available positions. OSVMLM application is built on Ionic, MongoDB, NodeJS and Gradient boosting machines (GBMs). OSVMLM is immensely useful for recruiters, as it makes data collection and processing effortless. This increases the visibility of various jobs open in different domains and suggests suitable job description to the job seekers. The key difference between OSVMLM and other platforms is that it is not limited to institutions and is accessible to all. This saves a lot of time for recruiters during initial screening. This proposed OSVMLM was validated with the real-time students dataset and achieved 91% prediction accuracy.

Index Terms— Gradient Boosting, Machine learning, Student Verification, Mobile Application, Classification

I. INTRODUCTION

In a developing country like India where crores of students graduate every year and with no point to collect data. It serves as medium that bridges this gap and sources for proof of credibility to verify details collected in the education phase. It aims to cluster according to their stream and send job alerts, fellowships and grants. It statistically improve the end on recruiter's end where processing job applications and reviewing candidates become easier and less time consuming. It aims to provide an integration of multiple services in different bundled into one application and single common platform that ensures all services are provided. It enables ease in integrating the

functionalities of multiple services in common platform. It is an open platform to be accessed by all. It follows a conceptual model that defines the structure and behaviour.

Job application portals are essential for the modern job market. They streamline the job search process for job seekers and employers alike. Job seekers can easily find and apply for jobs, while employers can post job openings and find qualified candidates (Borbye, 2008). Recent job application portals, such as Naukri, FreshersWorld, and Internshala, leverage innovative architectures and technologies to provide a seamless experience for users (Ng et al., 2023; Taherzadeh-Shalmai et al., 2023). Job application portals use Distributed architecture which distributes the system's components across multiple servers, making it scalable and resilient to failures. Microservices architecture breaks down the system into small, independent services that can be developed, deployed, and scaled independently, making it flexible and agile. Multi-tier architecture divides the system into different layers, such as the presentation layer, application layer, and database layer. Each layer has a specific responsibility, making the system easier to maintain and update (Al-Makhadmeh & Tolba, 2021). Job search portals use Cloud computing which allows job application portals to scale their infrastructure up or down as needed, making them more cost-effective and efficient to operate. Big data technology allows job application portals to analyze vast amounts of user data and job postings to provide personalized recommendations to job seekers and employers. Machine learning technology is used to develop algorithms that can automatically match job seekers with job postings, saving time and effort for both parties (Mahbub & Shubair, 2023; Zheng et al., 2023).

Machine learning (ML) is a rapidly evolving field with the potential to revolutionize the job market for both job seekers and job providers (Goldfarb et al., 2023). ML algorithms can be used to personalize the job search experience for job seekers by recommending jobs that match their skills, interests, and experience. This can save job seekers time and effort by helping them to focus on jobs that are likely to be a good fit for them. ML algorithms can also be used to screen resumes more efficiently and effectively, and to help job seekers prepare for interviews by providing them with feedback on their practice interviews and suggesting areas for improvement (Nanath & Olney, 2023). For job providers, ML algorithms can be used to source qualified candidates for job openings, to screen candidates more efficiently and effectively, and to help make more informed hiring decisions by providing insights into the candidates' skills, experience, and potential fit for the company. For example, LinkedIn uses ML to personalize the job search experience for its users, Indeed uses ML to match job seekers with job openings, Hiretual uses ML to help job providers screen candidates and make hiring decisions (Malik et al., 2023).

Integrating National Academic Depository (NAD), Instant Aadhar, and DigiLocker with job portals can offer several benefits to both job seekers and job providers (Tanwar et al., 2016). Job seekers can easily access and share their academic records and identity documents with job providers with just a few clicks, eliminating the need to collect and submit physical copies of documents. Job providers can verify the authenticity of job seekers' academic records and identity documents, reducing the risk of fraud. Job seekers can also have more control over who has access to their academic records and identity documents. Job providers can save time and resources by streamlining the process of collecting and verifying job seekers' documents. They can also reduce the costs associated with manual document processing and storage. Additionally, they can improve the security of job seekers' personal data by storing it in secure digital repositories (Kumar Jha et al., 2018; Kumar et al., 2020).

In this proposed system, we have developed an application that can bundle the services of all digital service websites and integrate them into one with some additional features. This application already has an in-built database that contains the Aadhar number and Phone Numbers of various members. Based on the profile details that the student will enter, the application will check whether the student's credentials and credibility are valid. Once the details are entered, the Aadhar number is matched with Aadhar number already present in our database. Once that is done, the student's fingerprint will be authenticated and verified using the fingerprint scanner on the smartphone. The main objective of this application is to offer services to the government and recruiters by giving them genuine and verified profiles.

Once a profile is created, the student can view and edit the profile if any discrepancy is found. On the administrator side, the data will be verified, grouped according to their stream, and given to the recruiter, who will be able to send out an alert for jobs available in different domains.

II. EXPERIMENTAL WORKFLOW ANALYSIS

The flow of this mobile application involves different modules, where at each module, the data is validated, and the desired result is provided for the user input. Firstly, the user is asked to either signup or log in according to their preference with the application. Once that process is done, the details are collected if signup has proceeded. Else, during login, the user can view the profile and edit if there are any changes. Following the signup process, the collected details will be name, email, age, phone number, college, department, Aadhar number, 10th %, 12th

%, CGPA, and fingerprint details. Once the pieces are verified using the built-in database in the application's backend, the new user becomes an existing user. The system architecture diagram proposes a basic workflow of this model where the user is taken into account from the front end, and the administrator handles the backend and maintenance of the application. We tried using a QR code to verify each profile; against each shape, there will be QR code to be scanned to verify this profile. The main disadvantage of this method is the X amount of time it needs to process and the Y amount of storage for each profile. It did not seem feasible and did not yield the output correctly. It took a lot of processing time, so we collected the biometric details and used the original factor authentication method to verify.

It makes use of Javascript and Android Studio, Ionic Libraries and JDK. It initially sets up where user sing=up and login page with existing account and user registration and verification is done and then profile updation and data medication takes place, finally receiving job alerts via e-mail will be done.

After the new user if completely signed and become an existing user, to improve the time factor of a recruiter while scrutiny. After grouping them according to their stream registration process will be done with rich data collection which can be used for government database and educational institutions for centralized database.

The user's details benefit an organization or the government if they want to send alerts regarding jobs, fellowships, and grants.

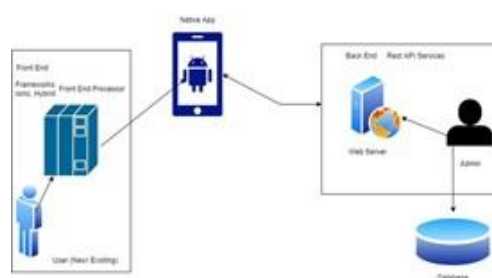


Figure 1. Workflow of proposed study

Dataset

Dataset, as we know it is a collection of data given to the system as a single unit for analytic and prediction purposes. The right kind of dataset is required to formulate the correct output of any application. In this application, we already have an in-built database containing the Aadhar numbers and Phone Numbers of the students in their respective institutions. The other dataset is students entering the details in the application, which will be collected in another database.

Initial Setup

First, we begin the initial setup process with a user signup and login page. Those already have an existing account log in to the application while new users go through the signup process.

User Registration and Verification

New users who want to avail of the application services have to sign- up for it. When you click signup, you are asked to enter your name, email id, and password. After that which, you are redirected to the application's detail gathering page where all the necessary student details and fingerprints are collected and, in the process, are verified using the Aadhar number and Phone Number already present in the backend database of the application. If those credentials do not match, the toast will display a message of invalidity. The collected details are stored in our application's database for different use cases.

Profile Updating and Data Modification

Once the new user has become an existing user, they can view the profile and check the entered details. They can edit their profile if any element is wrongly entered. In this updating process, the user can update if there are any new degrees acquired over a period and add that to their profile.

Receiving Job Alerts via Email

After the new user is wholly signed and becomes an existing user, they, along with all the current users, and grouped according to their stream. This is mainly done to improve the time factor of a recruiter while considering someone for a position. This same method can be applied to giving out alerts for fellowships, grants, and any other government-related job opportunities. After grouping them according to their stream, they receive a job alert via email through the email id they provided during the registration process. This rich data is used by the government

and several educational institutions that rely on google forms to collect their data during any placement training or an online assessment. Instead of collecting data every time that matches the company's requirements, they can collect the data once during the final year with an additional skill set. This is immensely helpful in filtering down the necessary candidates and sending them an alert for an assessment scheduled. This will save a lot of time, and students don't just re-enter details every time a company arrives for a placement. It is hassle-free and easier to use.

III. METHODOLOGY

Implementing this application in real-time will prove to be successful even with dummy datasets that were created. This application's prototype passed all testing processes and will be helpful for any recruiter of an organization. Five major job portals in our country carry to have one major flaw. That flaw is verification since they don't have access to the government's database containing the Aadhar numbers of the graduates. This creates a problem of junk profiles. Many junk profiles can be completed in these job portals, making it challenging to organize and categorize the data. This drawback prevents them from providing the correct Number of jobs available to the particular domain. Even though they have an option of uploading a resume and all those other extra features, not being able to provide good profiles places it a notch below. Our country's five major job portals are Naukri.com, Monster, Shine.com, Indeed, and Jobs for her. All these platforms have all the primary input fields contained in our application. One by one, we will analyze the use of websites and how their verification works. Having an account in each of these portals, we will dissect the differences between our application and the job portals and plot a basic comparison graph containing all the features of our verification, input fields, and response time.

Naukri.com

Fronting one of the biggest job portals, its registration process is quite simple and contains almost all the input fields such as name, email id, setting a password for the account, phone number, and an option to upload their resume. It does not mention collecting Aadhar number or any biometric detail. They have a peak response time.

MonsteIndia.com

MonsterIndia.com's job portal is quite interesting as they have added a couple more input fields than our application. This portal turns out to be collecting all the essential skills and current location, resume, total years of work experience, preferred industry, and preferred department.



Figure 2. Platform comparison bar chart

It also does not account for collecting Aadhar numbers and biometric details. It has an average response time.

Jobs for her

Jobs for her portal mainly focuses on providing women employment and collecting only their educational details. Though not open to all their input fields, it is worth looking for a comparison for our application. They also use standard input fields such as email id and phone number. Their profile registration process is quite lengthy, containing different sections. Even with a tedious registration process, they do not collect Aadhar numbers and biometric details. They have an average response time as well.

Shine.com

Also, one of India's largest job portals has quite a similar registration process of collecting phone numbers, email id and it follows a sectional registration process like Jobs for her, with the input fields varying from locations, preference, uploading a resume, and even having a separate column for employer's registration. This employer column makes this a unique portal because this feature is not really available on all the other platforms. With all

these fantastic features and a lengthy registration process, it does not contain fields to collect Aadhar numbers or any biometric detail. They have an average response time.

Indeed

Indeed, it is one of those platforms where you only need an email id to enter the portal and look for a job. It offers no sort of verification except activating email id. Their features consist of uploading a resume. Once the email id is entered, you can browse through the web application, create your profile, and strengthen it. One of the features of this portal is that it offers a salary guide where the filter can push for a job according to the salary per year. This feature turns out to be a value add for people scrubbing through for employment who have experience.

Comparison Graph

A comparison chart, in a nutshell, is a visual depiction of numerous possibilities that may be compared to one another. It might include their advantages, disadvantages, essential characteristics, and so on, all of which would highlight significant distinctions between these entities (or choices).

Fig. 2 represents a bar graph comparison chart between the application and the country's other five major job portals. The key difference that makes the application stand out is verifying Aadhar numbers and collecting biometric details, which ultimately help the recruiter in the initial screening process.

Integrated testing for group of components, Table 1. Test cases are described with expected output and actual output. Test Case 1 depicts user registration, entering credentials for Test Case 2, viewing profile in Test Case 3, Test case 4 illustrates receiving jobs alerts via email. All are evaluated according to the standards and for interactive framework with more security.

TABLE I. TEST CASES

INPUT	EXPECTED OUTPUT	ACTUAL OUTPUT	RESULT
Test Case 1: User Registration			
Entering valid username and password	Display the page to enter details	Page to enter the details was displayed	Passed
Entering invalid username and password	Display an error message	An error message is displayed	Passed
Test Case 2: Entering Credentials			
Entering necessary details of the graduate	Profile creation and viewing it	Profile was created successfully	Passed
Entering the invalid details of the graduate	A toast message of invalid credentials	Invalid credentials message displayed	Passed
Test Case 3: Viewing Profile			
Entering valid username and password	Displays the graduate's profile	Graduate's profile displayed	Passed
Entering invalid username and password	Displays an error message	An error message is displayed	Passed
Test Case 4: Receiving Job Alerts via Email			
Send Job Alert in the job alerts tab	A job alert to be sent to all the graduates present in the database	Job Alerts were sent to all registered graduates.	Passed

IV. RESULTS

The One Point Student Verification Machine Learning Model (OSVMLM) is a powerful tool designed to streamline the verification and recruitment process for both students and recruiters. The OSVMLM has successfully addressed the challenges associated with tracking and processing data for millions of graduating students each year. By leveraging advanced machine learning techniques and a robust technology stack, it has revolutionized student verification and job placement in the following ways:

Comprehensive Data Collection: OSVMLM collects extensive data on students, including biometric information, secondary education details, a list of completed degrees, and detailed information on their current or last degree. This rich dataset enables accurate and thorough student profiling.

Biometric Verification: The system incorporates biometric data, including fingerprints, to ensure the authenticity of student information. The integration with the government's Aadhar database further enhances data accuracy and trustworthiness.

Automated Ranking and Job Matching: OSV MLM assigns ranks to student profiles based on their qualifications and skills. It autonomously matches student profiles with current job descriptions, simplifying the job search process for students.

Effortless Data Processing: Recruiters benefit from OSV MLM's seamless data collection and processing capabilities. It significantly reduces the effort required for initial screening and candidate selection.

Increased Job Visibility: The system enhances job visibility across various domains, allowing students to identify and apply for available positions more easily.

Inclusivity: Unlike some other platforms that are limited to specific institutions, OSV MLM is accessible to all students, broadening opportunities and simplifying the recruitment process.

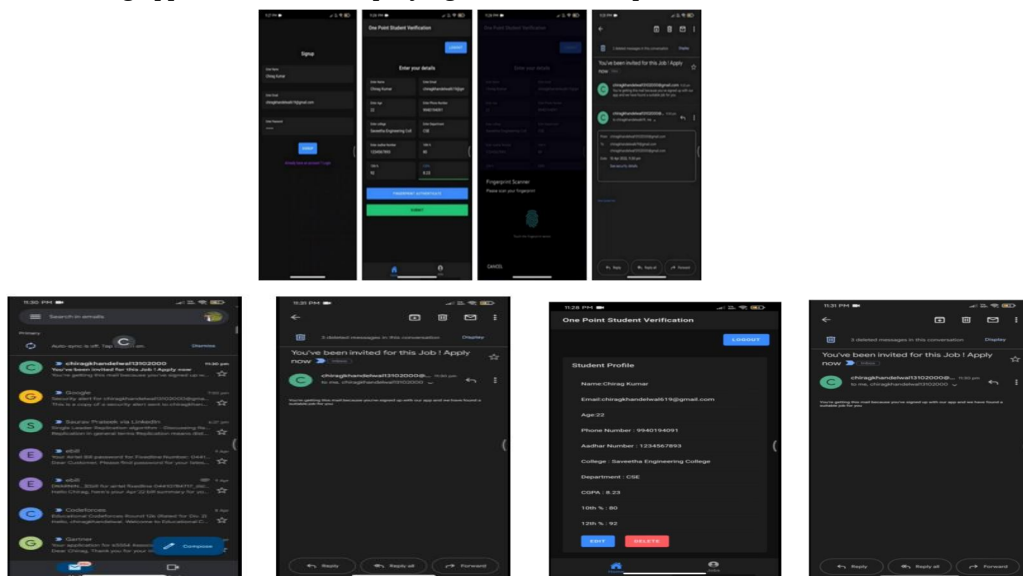


Figure 3. Output screenshots of the developed mobile application

The above figures represent the following process. This is the login and signup page. A new user will be able to sign up for our platform and will be able to register their details along with their biometric information. Existing users can view their profiles, edit, and modify details if there are discrepancies. To log in, the user needs to enter his/her email id and password and to signup to enter their details. The process for signup is straightforward. The user needs to enter his/her name and email-id and create a new password to register with the platform. Once done, the application will show the form to collect their educational details and the Aadhar number. A typical database is present in the backend with the details of Aadhar number and Phone Number, which will serve as a primary key. The verification process goes through only if the Aadhar number in the common database matches with the Aadhar number entered else. A toast message of Invalid Credentials is displayed. Once the sign up process is completed, the user can log in and edit his profile if there are any discrepancies in the details. The biometric detail stored is the fingerprint which is authenticated during the signup process.

In comparing our application with five major job portals in India, namely Naukri.com, MonsterIndia.com, Jobs for her, Shine.com, and Indeed, several key differences and similarities emerge.

Naukri.com, being one of the largest job portals, offers a simple registration process with essential input fields such as name, email id, password, and phone number. However, it lacks the collection of Aadhar numbers or biometric details, which could impact data accuracy. Its response time is noted to be peak.

MonsterIndia.com, on the other hand, expands on the input fields by including additional details such as skills, current location, and years of work experience. While it provides more comprehensive information, it also lacks Aadhar number collection and has an average response time.

Jobs for her focuses primarily on providing employment opportunities for women and collects only educational details during registration. While its registration process is lengthy, it does not include Aadhar number or biometric detail collection, similar to the other portals with average response time.

Shine.com follows a sectional registration process similar to Jobs for her, with additional features like a separate section for employer registration. However, like the others, it does not collect Aadhar numbers or biometric details, despite offering a variety of input fields. Its response time is average as well.

Indeed stands out for its simplicity, requiring only an email id for registration and activation. While it offers features like resume uploading and a salary guide, it also lacks Aadhar number or biometric detail collection, making it similar to the other portals in terms of data verification.

Overall, while these job portals offer various features and input fields, they share the common limitation of not collecting Aadhar numbers or biometric details for verification purposes, unlike our application. This distinction could significantly impact the accuracy and reliability of the data collected and processed by these platforms compared to our comprehensive verification approach.

V. CONCLUSION AND FUTURE ENHANCEMENT

The objective of the application is to collect student details who have acquired the degrees and to verify through our platform, which creates the overall credibility of the profile. Students are also filtered according to their stream of education. This improves the scope of the recruiter of an organization that has sorted and filtered data. The student can apply for many available jobs through one common platform. This platform serves as an integration of Instant Aadhar, National Academic Depository, and DigiLocker. This is a common platform for including student details, and the recruiter presents them with many jobs in every location. A GPS can be added that shows the positions available in their present area first and then moves to other sites. Also, students can be requested to add a resume for the ease of the recruiter.

REFERENCES

- [1] Al-Makhadmeh, Z., & Tolba, A. (2021). Independent and tailored network-slicing architecture for leveraging industrial internet of things job processing. *Computer Networks*, 187, 107827. <https://doi.org/10.1016/J.COMNET.2021.107827>
- [2] Borbye, L. (2008). Professional conduct during the job application process. *Secrets to Success in Industry Careers*, 27–42. <https://doi.org/10.1016/B978-012373869-1.50004-2>
- [3] Goldfarb, A., Taska, B., & Teodoridis, F. (2023). Could machine learning be a general purpose technology? A comparison of emerging technologies using data from online job postings. *Research Policy*, 52(1), 104653. <https://doi.org/10.1016/J.RESPOL.2022.104653>
- [4] Kumar Jha, R., Srivastav, Y., Sumbli, V., Trisha, Gandhi, V., & Jain, S. (2018). RFID based food rationing system. *HardwareX*, 4, e00043. <https://doi.org/10.1016/J.OHX.2018.E00043>
- [5] Kumar, R., Khayyam, K. U., Singla, N., Anand, T., Nagaraja, S. B., Sagili, K. D., & Sarin, R. (2020). Nikshay Poshan Yojana (NPY) for tuberculosis patients: Early implementation challenges in Delhi, India. *Indian Journal of Tuberculosis*, 67(2), 231–237. <https://doi.org/10.1016/J.IJTb.2020.02.006>
- [6] Mahbub, M., & Shubair, R. M. (2023). Contemporary advances in multi-access edge computing: A survey of fundamentals, architecture, technologies, deployment cases, security, challenges, and directions. *Journal of Network and Computer Applications*, 219, 103726. <https://doi.org/10.1016/J.JNCA.2023.103726>
- [7] Malik, A., Onyema, E. M., Dalal, S., Lilhore, U. K., Anand, D., Sharma, A., & Simaiya, S. (2023). Forecasting students' adaptability in online entrepreneurship education using modified ensemble machine learning model. *Array*, 19, 100303. <https://doi.org/10.1016/J.ARRAY.2023.100303>
- [8] Nanath, K., & Olney, L. (2023). An investigation of crowdsourcing methods in enhancing the machine learning approach for detecting online recruitment fraud. *International Journal of Information Management Data Insights*, 3(1), 100167. <https://doi.org/10.1016/J.JJIMEI.2023.100167>
- [9] Ng, J. W. X., Song, Z., & Lievens, F. (2023). Human capital effects in the job search process for new labor market entrants: A double-edged sword? *Journal of Vocational Behavior*, 144, 103894. <https://doi.org/10.1016/J.JVB.2023.103894>