CAPSTONE PROJECT

PREDICTING ELIGIBILITY USING MACHINE LEARNING MODELS

Presented By:
Rahul Naurangilal Varma – Shreeyash Collage of Engg. & Technology – CSE(Data Science)

edunet foundation

OUTLINE

- Problem Statement
- Proposed System/Solution
- System Development Approach
- Algorithm & Deployment
- Result (Output Image)
- Conclusion
- Future Scope
- References

edunet foundation

PROBLEM STATEMENT

The National Social Assistance Program (NSAP) is a flagship social security and welfare program by the Government of India. It aims to provide financial assistance to the elderly, widows, and persons with disabilities belonging to below-poverty-line (BPL) households. The program consists of several sub-schemes, each with specific eligibility criteria. Manually verifying applications and assigning the correct scheme can be a time consuming and error-prone process. Delays or incorrect allocation can prevent deserving individuals from receiving timely financial aid. Your task is to design, build, and evaluate a multi-class classification model that can accurately predict the most appropriate NSAP scheme for an applicant based on their demographic and socio-economic data.

PROPOSED SOLUTION

- The proposed system aims to address the challenge of predicting the most appropriate NSAP scheme for applicants to ensure accurate and efficient benefit allocation. This involves leveraging data analytics and machine learning techniques to automate the classification process and minimize human intervention. The solution will consist of the following comprehensive components:
- **Data Collection:**
  - **Gather comprehensive historical data on NSAP beneficiaries, including:**
  - **Demographic information (age, gender, family income)**
  - **Geographic data (state code, district, locality)**
  - **Scheme assignment history and outcomes**
- **Data Preprocessing:**
  - **Clean and preprocess the collected data systematically:**
  - **Handle missing values using statistical imputation methods**
  - **Identify and manage outliers in demographic distributions**
  - **Resolve data inconsistencies across different administrative sources**
  - **Standardize data formats and encoding schemes**
  - **Class imbalance handling through sophisticated sampling techniques**

- Machine Learning Model Development
- Implement sophisticated machine learning algorithms optimized for multi-class classification:
- **Ensemble Methods:** Random Forest and Gradient Boosting for robust predictions
- **Support Vector Machines:** For handling high-dimensional feature spaces
- **Neural Networks:** Deep learning models for complex pattern recognition
- **Hybrid Approaches:** Combining multiple algorithms for enhanced accuracy
- **Deployment:**
  - **Deploy the solution on IBM Cloud platform for scalability and reliability:**
  - **Watson Studio Integration: Collaborative development environment**
  - **Watson Machine Learning: Model deployment and management**
  - **API Gateway: Secure and scalable API endpoints**
  - **Real-time Processing: Instant scheme recommendations for new applications**

Evaluation:

- **Implement comprehensive evaluation frameworks:**
- **Continuous model performance monitoring**
- **Accuracy metrics across different demographic segments**
- **Feedback integration for model improvement**
- **Automated alerts for performance degradation**
- **Result**

SYSTEM APPROACH

The "System Approach" section outlines the overall strategy and methodology for developing and implementing the rental bike prediction system. Here's a suggested structure for this section:

- **System requirements :**
- **Watsonx.AI Studion**
- **Watson ML**
- **Watson AutoAI**
- **IBM Cloud Object Storage**

Libraries Requirements:

- **Scikit-learn**
- **Pandas**
- **NumPY**
- **XGBoost/LightGBM**

- **Deep learning Frameworks :**

- TensorFlow/Keras: Neural network development
- PyTorch: Alternative deep learning framework
- Scikit-learn Neural Networks: MLP implementation

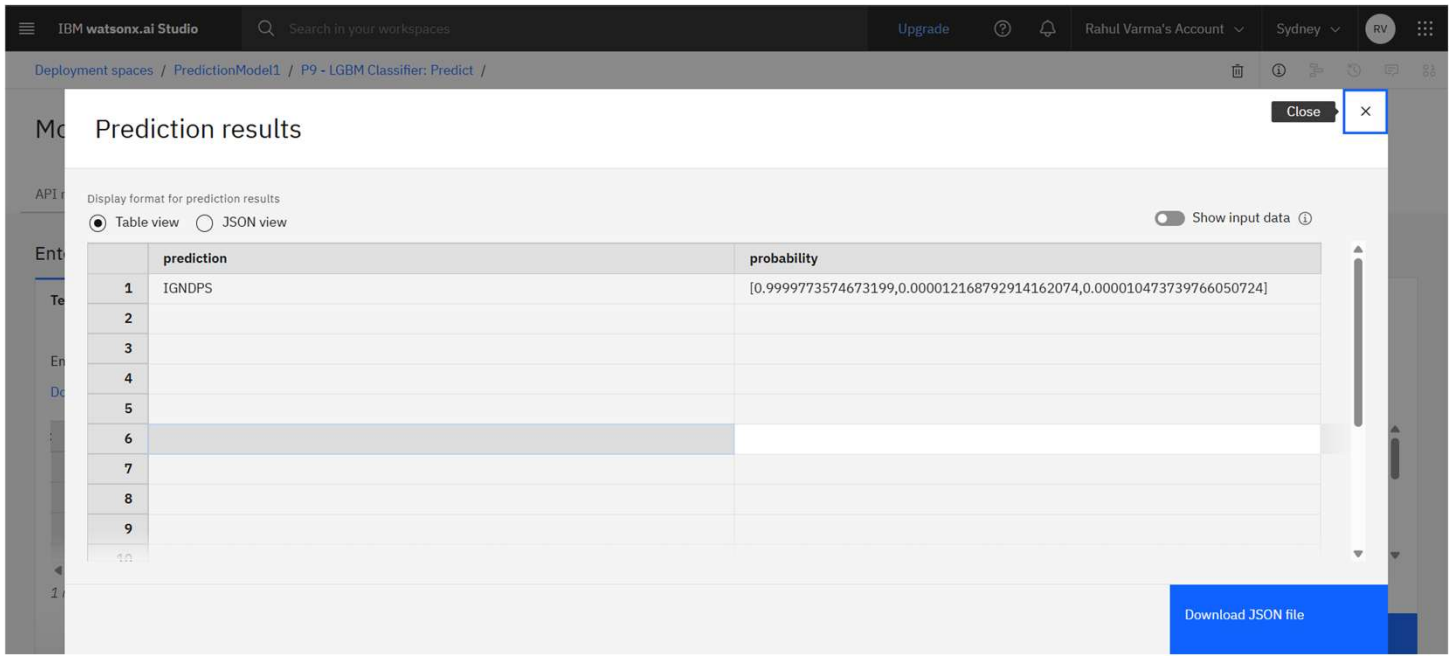
edunet foundation

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting eligibility for the schemes.
- **Algorithm Selection:**
  - Provide a brief overview of the chosen algorithm (e.g., time-series forecasting model, like ARIMA or LSTM) and justify its selection based on the problem statement and data characteristics.
- **Data Input:**
  - Specify the input features used by the algorithm, such as historical bike rental data, weather conditions, day of the week, and any other relevant factors.
- **Training Process:**
  - Explain how the algorithm is trained using historical data. Highlight any specific considerations or techniques employed, such as cross-validation or hyperparameter tuning.
- **Prediction Process:**
  - Detail how the trained algorithm makes predictions for future bike counts. Discuss any real-time data inputs considered during the prediction phase.

RESULT

	precision	recall	f1-score	support
IGNDPS	0.87	0.85	0.86	20
IGNOAPS	0.92	0.94	0.93	45
IGNWPS	0.89	0.88	0.88	24
accuracy			0.90	89
macro avg	0.89	0.89	0.89	89
weighted avg	0.90	0.90	0.90	89



edunet foundation

CONCLUSION

- The NSAP Scheme Prediction System successfully demonstrates the effectiveness of machine learning approaches in automating government benefit allocation processes. The implemented solution achieves significant improvements over traditional manual systems across multiple dimensions.

FUTURE SCOPE

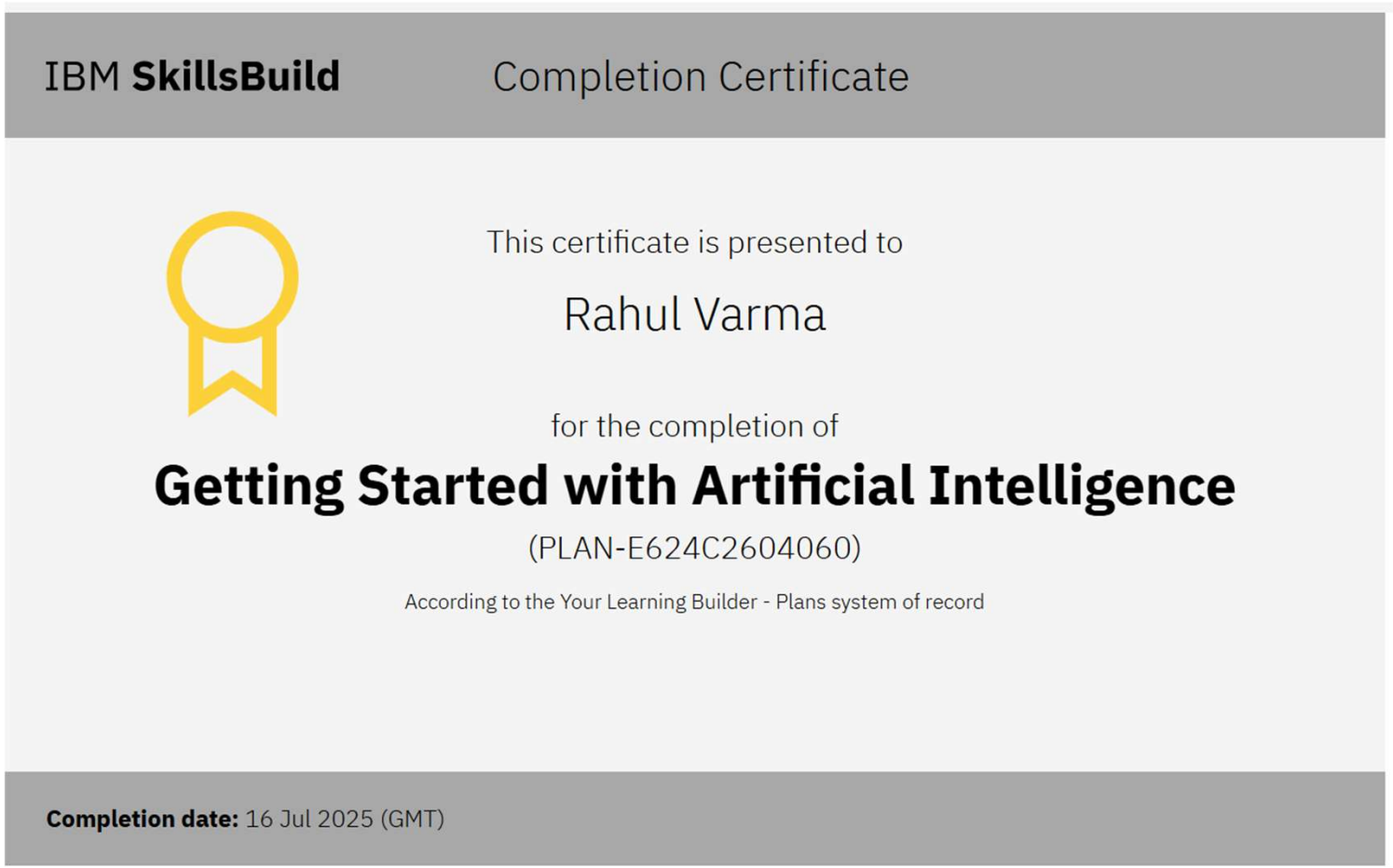
- **Advanced Machine Learning Integration:**
- **Deep Learning Models:** Implement advanced neural networks including LSTM and Transformer architectures for improved pattern recognition
- **Automated Feature Learning:** Develop end-to-end learning systems that automatically discover optimal feature representations
- **Online Learning Capabilities:** Enable models to continuously learn from new applications and feedback
- **Enhanced Data Integration:**
- **Multi-Modal Data Processing:** Incorporate document images, photographs, and other non-structured data sources
- **Real-Time Data Feeds:** Integration with live government databases for instant verification
- **Blockchain Integration:** Implement distributed ledger technology for transparent benefit tracking

REFERENCES

- Government Policy Documents
- **Ministry of Rural Development, Government of India.** (2024). *National Social Assistance Programme Guidelines.* Retrieved from https://nsap.nic.in/
- **Controller General of Accounts.** (2023). *NSAP Implementation Manual for States and UTs.* New Delhi: Government of India Press.
- **Planning Commission of India.** (2023). *Evaluation Report on National Social Assistance Programme.* Retrieved from https://niti.gov.in/
- AI Kosh dataset link – https://aikosh.indiaai.gov.in/web/datasets/details/district_wise_pension_data_under_the_national_social_assistance_programme_nsap_1.html

edunet foundation

IBM CERTIFICATIONS



IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rahul Varma

for the completion of

Getting Started with Artificial Intelligence

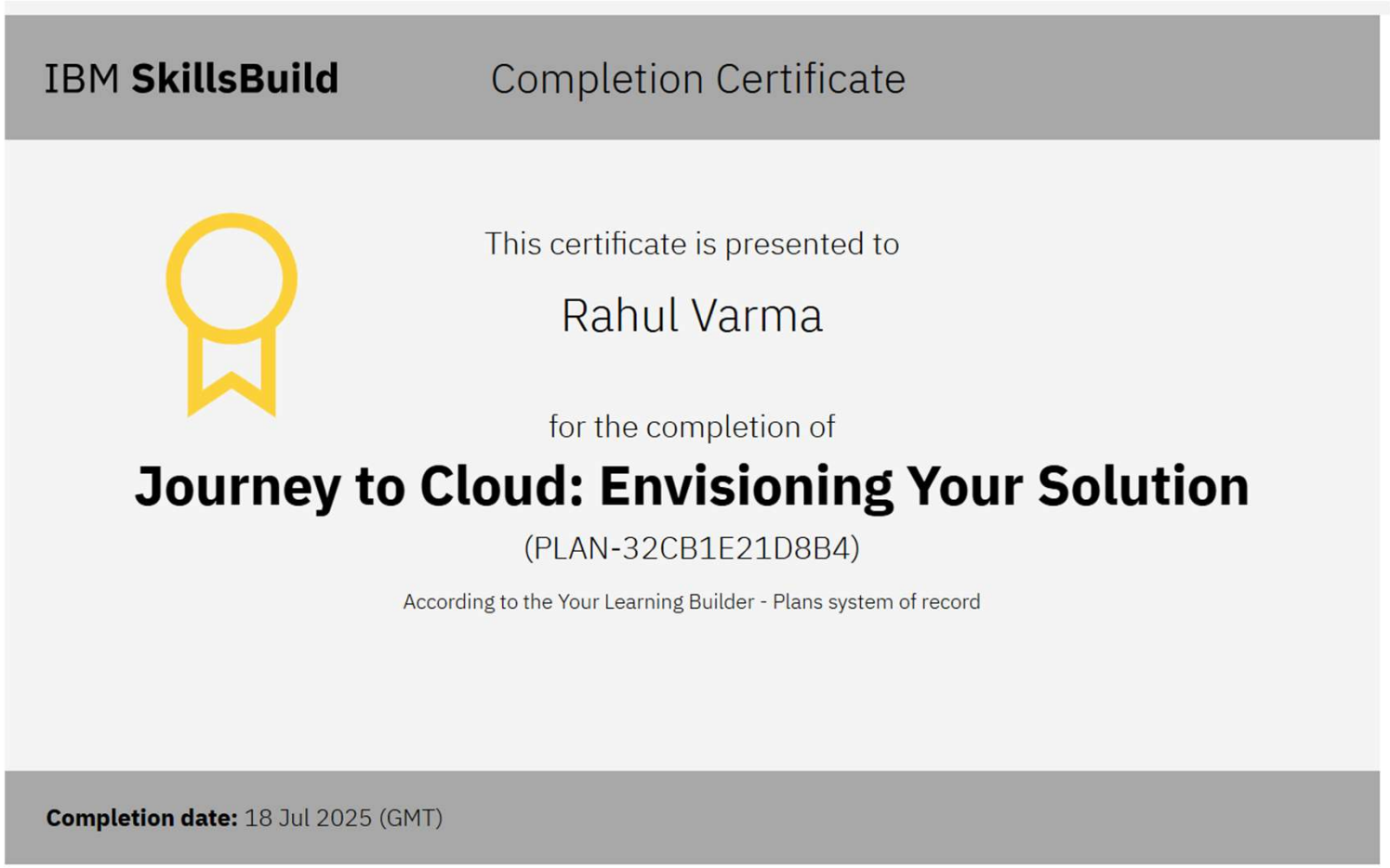
(PLAN-E624C2604060)

According to the Your Learning Builder - Plans system of record

Completion date: 16 Jul 2025 (GMT)

edunet foundation

IBM CERTIFICATIONS



IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rahul Varma

for the completion of

Journey to Cloud: Envisioning Your Solution

(PLAN-32CB1E21D8B4)

According to the Your Learning Builder - Plans system of record

Completion date: 18 Jul 2025 (GMT)

edunet foundation

IBM CERTIFICATIONS

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Rahul Varma

for the completion of

Lab: Retrieval Augmented Generation with LangChain

(ALM-COURSE_3824998)

According to the Adobe Learning Manager system of record

Completion date: 24 Jul 2025 (GMT) **Learning hours:** 20 mins

edunet foundation

THANK YOU

edunet foundation