CAPSTONE PROJECT

PENSION ELIGIBILITY CLASSIFIER

Presented By:
1. Subin Kumar- Lakshmi Narain College of Technology, Bhopal-Computer Science and Engineering

edunet foundation

OUTLINE

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

Problem Statement

Problem statement 34 - Predicting Eligibility using Machine Learning

The Challenge:
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. The goal is to create a reliable tool that could assist government agencies in quickly and accurately categorizing applicants, ensuring that benefits are delivered to the right people efficiently.

Proposed Solution

The system predicts the correct NSAP scheme for applicants using machine learning. This helps deliver financial aid faster and more accurately.

- **Data Collection**
  - Used AI Kosh NSAP dataset.
  - Features: Age, gender, income, disability, marital status.
  - Labels: Scheme codes (IGNDPS, IGNOAPS, IGNWPS).

- **Data Preprocessing**
  - Cleaned missing or inconsistent data.
  - Encoded categorical fields into numbers.

- **Machine Learning Model**
  - Applied Random Forest Classifier in IBM Watsonx.ai.
  - Trained to classify applications into the correct scheme.

Proposed Solution

- **Deployment**
  - Deployed as an online model in Watsonx.ai.
  - Allows instant predictions on new applicant data.

- **Evaluation**
  - Model outputs prediction confidence (50–70%).
  - Example results:
    - IGNDPS – 70%
    - IGNOAPS – 60%
    - IGNWPS – 70%
- **Result**
  - The solution improves accuracy and saves time in scheme allocation.
  - Clear prediction scores help in decision-making.

System Approach

The strategy, tools, and resources used to build and deploy the NSAP eligibility prediction system.

- **System Requirements**
  - IBM Cloud Lite Account
  - IBM Watsonx.ai Studio for model training and deployment
  - Web browser for accessing the platform
- **Libraries and Tools**
  - Watson Machine Learning for creating and managing models
  - AutoAI or Random Forest classifier module
  - Data Refinery for data cleaning and preparation
  - JSON export for prediction results

Algorithm & Deployment

- **Algorithm Selection:**
  - I used a Random Forest Classifier because it handles categorical and numerical data well and works effectively for multi-class classification tasks.

- **Data Input:**
  - The model used these features such as State and district codes, Number of beneficiaries by gender, Category-wise counts (General, OBC), Aadhaar and mobile number counts, Financial year etc.
  - The target variable was schemecode, representing the NSAP scheme.

- **Training Process:**
  - Trained on 2157 records in Watsonx.ai Studio.
  - AutoAI handled data splitting and encoding of categorical fields.
  - The model was evaluated with prediction confidence scores.

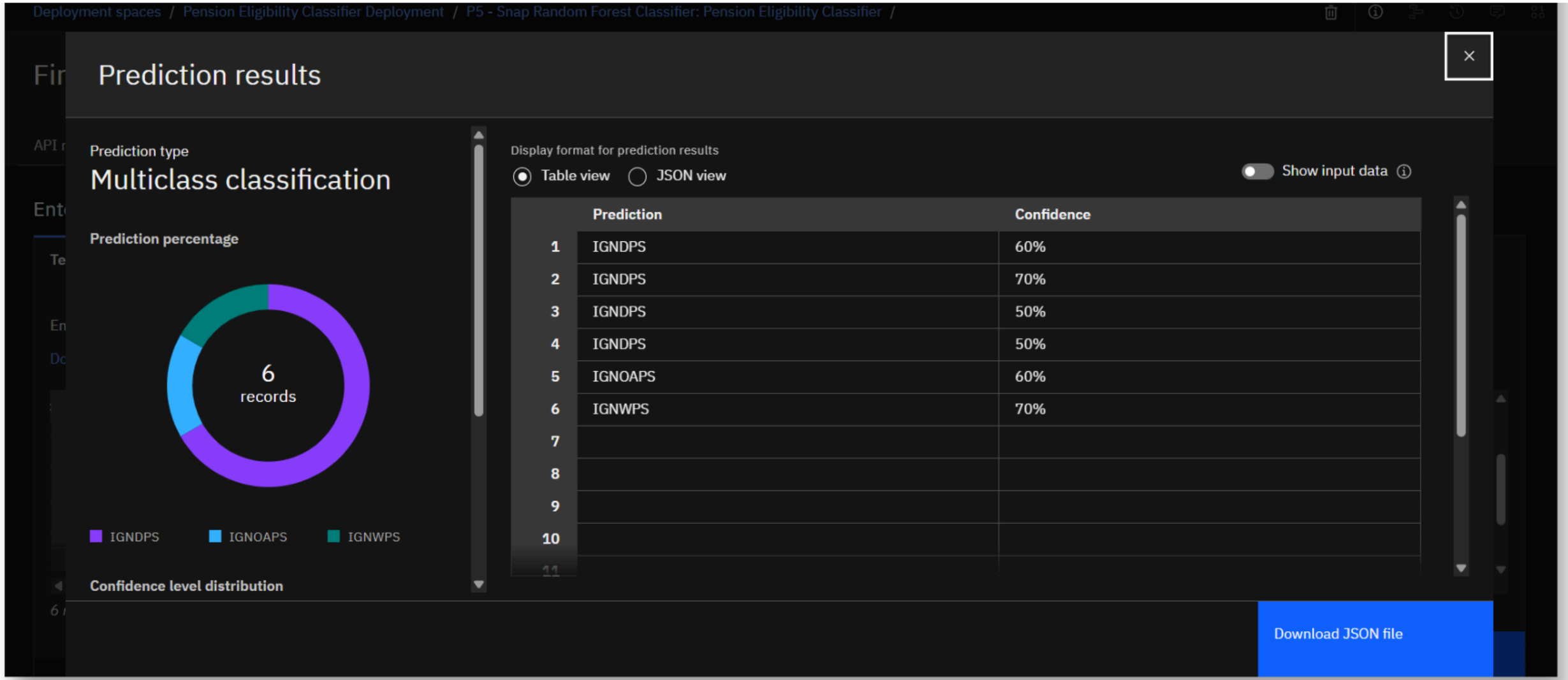
- **Prediction Process:**
  - New applicant data is input via Watsonx.ai.
  - The model predicts the most likely NSAP scheme and provides a confidence percentage (e.g., 60–70%).
  - Results can be exported in JSON format for record-keeping.

Result

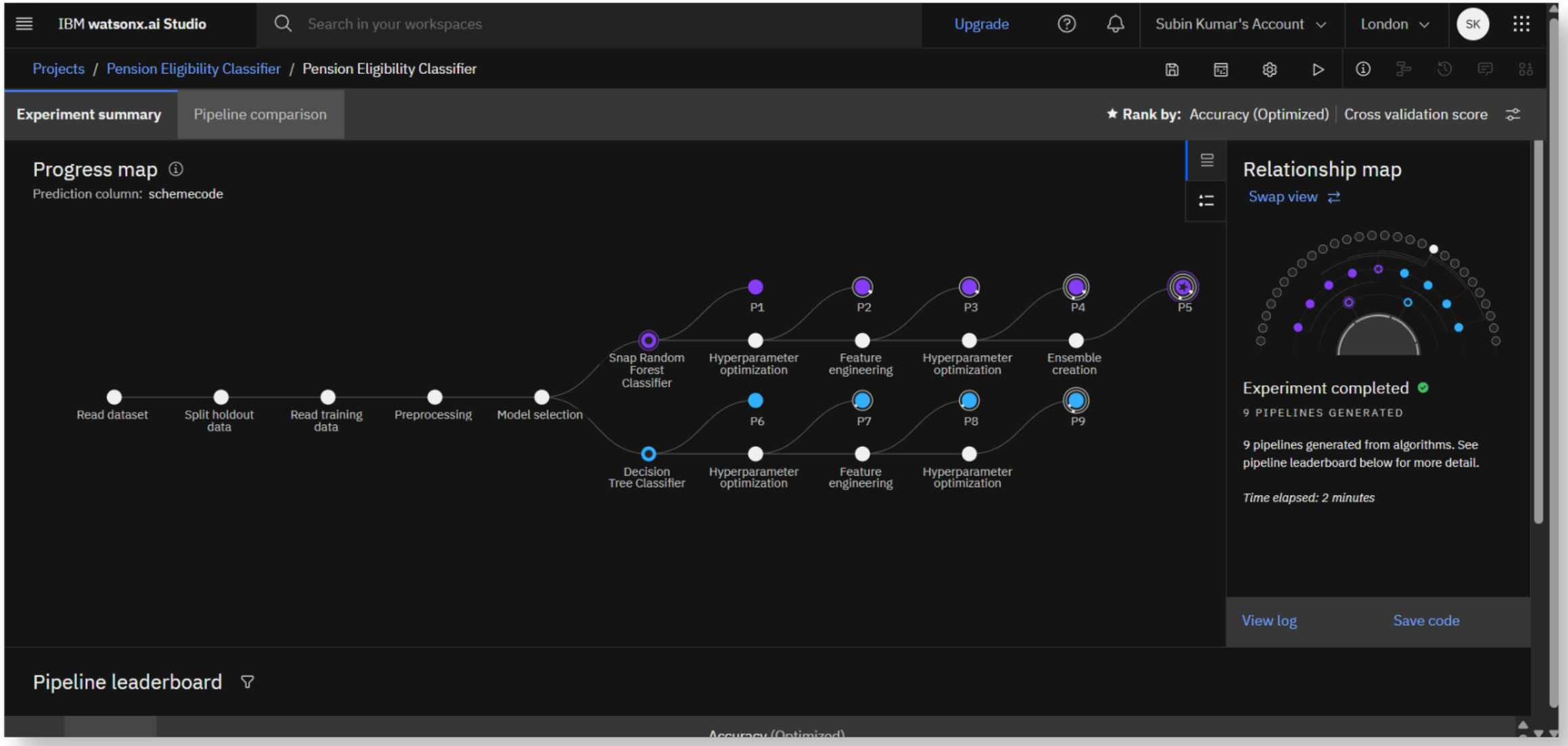
- The Random Forest Classifier predicts the eligible NSAP scheme with confidence scores ranging from 50% to 70% on test records.
- Example predictions:
  - IGNDPS (70%)
  - IGNOAPS (60%)
  - IGNWPS (70%)
- The model was deployed successfully in IBM Watsonx.ai and tested with real data inputs.
- This solution improves classification speed and reduces manual workload.

GitHub Repo Link: https://github.com/Subinkumar077/pension-eligibility-classifier

Result


Deployment spaces / Pension Eligibility Classifier Deployment / P5 - Snap Random Forest Classifier: Pension Eligibility Classifier /
Prediction results
Prediction type
Multiclass classification
Prediction percentage
6
records
IGNDPS
IGNOAPS
IGNWPS
Confidence level distribution
Display format for prediction results
Table view
JSON view
Show input data
Prediction
Confidence
1 IGNDPS 60%
2 IGNDPS 70%
3 IGNDPS 50%
4 IGNDPS 50%
5 IGNOAPS 60%
6 IGNWPS 70%
7
8
9
10
Download JSON file

Result



Result

Pipeline leaderboard

	Rank	Name	Algorithm	Specialization	Accuracy (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 5	Batched Tree Ensemble Classifier (Snap Random Forest Classifier)	INCR	0.984	HPO-1 FE HPO-2 BATCH	00:00:39
	2	Pipeline 4	Snap Random Forest Classifier		0.984	HPO-1 FE HPO-2	00:00:37
	3	Pipeline 3	Snap Random Forest Classifier		0.984	HPO-1 FE	00:00:29

Deployments | Model details

Search | New deployment

Name	Type	Status	Tags	Last modified
Final Deployment Pension Eligibility Classifier	Online	Deployed		23 seconds ago Subin Kumar (You)

About this asset

Name
P5 - Snap Random Forest Classifier: Pension Eligibility Classifier

Description
No description provided.

Asset Details

edunet foundation

Conclusion

- The proposed machine learning model successfully predicts the appropriate NSAP scheme for applicants using socio-economic and demographic data.

- The deployment in IBM Watsonx.ai enables faster and more accurate classification, reducing manual verification time.

- **Challenges faced:**
  - Handling missing or inconsistent data entries
  - Encoding multiple categorical fields correctly

- **Future improvements:**
  - Adding more recent data to improve model accuracy
  - Exploring other algorithms for better performance

- Accurate prediction of eligibility ensures that financial assistance reaches the right people efficiently and on time.

Future scope

- **Incorporate more data:** Include real-time applicant-level data such as age, income proofs, disability certificates, etc., to improve prediction accuracy.
- **Model optimization:** Experiment with advanced algorithms like XGBoost or LightGBM for better performance.
- **Regional expansion:** Extend the system to cover more districts and states across India, supporting localized scheme allocation.
- **API Integration:** Develop APIs to connect the model with government application portals for live eligibility checks.
- **Edge deployment**: In the future, run lightweight versions of the model on local devices (e.g., district offices) for offline prediction.
- **Multilingual UI:** Build user-friendly, regional-language interfaces to help field workers use the system more easily.

References

- AI Kosh Dataset – [District-wise Pension Data under NSAP (National Social Assistance Programme)]
  - https://aikosh.indiaai.gov.in/web/datasets/details/district_wise_pension_data_under_the_national_social_assistance_programme_nsap_1.html
- IBM Watsonx.ai Documentation –
  - https://dataplatform.cloud.ibm.com/docs/content/wsj/getting-started/get-started-wdp.html?context=cpdaas&audience=wdp
- YouTube – IBM Watsonx.ai tutorials & model deployment guides
  - https://youtu.be/swPBNKKPK0E?si=xIWi-3urTKi8Jg2j
  - https://youtu.be/3sav6vUG_XQ?si=WozswG0VEERrEbIR

IBM Certifications

In recognition of the commitment to achieve professional excellence

Getting Started with Artificial Intelligence
IBM SkillsBuild
IBM

Subin Kumar

Has successfully satisfied the requirements for:

Getting Started with Artificial Intelligence

Issued on: Jul 16, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/6f64cda1-cf6d-483e-99e3-1c5f5f529bd1

IBM

edunet foundation

IBM Certifications

In recognition of the commitment to achieve professional excellence

Journey to Cloud: Envisioning Your Solution
IBM SkillsBuild
IBM

Subin Kumar

Has successfully satisfied the requirements for:

Journey to Cloud: Envisioning Your Solution

Issued on: Jul 19, 2025
Issued by: IBM SkillsBuild

Verify: https://www.credly.com/badges/8562a5aa-be4b-463d-999a-1bb058dfcba8

IBM

edunet foundation

IBM Certifications

IBM **SkillsBuild** Completion Certificate

This certificate is presented to

Subin Kumar

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

THANK YOU

edunet
foundation