CAPSTONE PROJECT

ANNUAL DEATHS BY AGES (AUTO-AI)

Presented By:

1. **Medisetti Navya Sri**
   Pragati Engineering College-ECE

edunet foundation

OUTLINE:

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

PROBLEM STATEMENT:

- Analyze and understand the trends and patterns of annual deaths by age group across the world from1950 to 2023.
- Examine the factors contributing to increasing death rates among adults aged 60-79.
- Develop effective strategies and interventions to address the disparities in mortality rates across age groups across the world.

PROPOSED SOLUTION

- **The solution will consist of the following components:**
- **Data Collection:**

  -Gather the required information regarding to different countries , year(1950-2023), death: sex: estimations.

- **Data Pre-processing:**

  -Feature engineering to extract relevant features from the data that might impact annual deaths across the world.

- **Machine Learning Algorithm:**

  -Implement a machine learning algorithm, such as a time-series forecasting model (e.g., ARIMA, SARIMA, or LSTM), to predict annual death rates by different ages based on historical patterns.

  -Consider incorporating other factors like entity ,year, death sex: estimate : varaints are events to improve prediction accuracy.

- **Deployment:**

  -Develop a user-friendly interface or application that provides predictions for annual death by ages across the world.

- **Evaluation:**

  -Assess the model's performance using appropriate metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Median absolute error (MedAE) or other relevant metrics. continuous monitoring of prediction accuracy.

SYSTEM APPROACH

This section consists of analyzing annual death by age using watson studio machine learning:

- **System components:**

Data Ingestion, Data preparation, Exploratory Data Analysis, Machine learning model development, Model depolyment, Model monitoring and maintenance.

- **System benefits:**

Improved accuracy, Enhanced insights, Informed decision-making, Scalability and others like collaboration.

ALGORITHM & DEPLOYMENT

- In the Algorithm section, describe the machine learning algorithm chosen for predicting annual death rate by ages. Here's an example...
- **Algorithm Selection:**

In chosen machine learning for annual deaths by ages mainly used algorithm are snap random forest regressor and ridge regressor.

➢ **Random Forest Regressor:**

-An ensemble learning method that combines multiple decision tree to predict continuous outcomes.

-It is worked by bootstrapping samples from the training data.

-Training a decision tree on each sample. Averaging the predictions from all trees.

-It is used to when dealing with the complex, non-linear relationships between features and target variables.

ADVANTAGES:

- Handles ,high-dimensional data, robust to outliers and interpretable.

Experiment summary | Pipeline comparison | ★ Rank by: Root mean squared error (RMSE) (... | Cross validatic

	Rank	Name	Algorithm	RMSE (Optimized) Cross Validation	Enhancements	Build time
★	1	Pipeline 4	Snap Boosting Machine Regressor	601.957	HPO-1 FE HPO-2	00:01:07
	2	Pipeline 3	Snap Boosting Machine Regressor	637.504	HPO-1 FE	00:00:53
	3	Pipeline 2	Snap Boosting Machine Regressor	687.081	HPO-1	00:00:08
	4	Pipeline 8	Snap Random Forest Regressor	758.833	HPO-1 FE HPO-2	00:01:12
	5	Pipeline 7	Snap Random Forest Regressor	758.833	HPO-1 FE	00:00:40
	8	Pipeline 5	Ridge	973.572	None	00:00:02
	7	Pipeline 6	Ridge	973.530	HPO-1	00:00:07
	4	Pipeline 7	Ridge	866.724	HPO-1 FE	00:00:34
	3	Pipeline 8	Ridge	866.724	HPO-1 FE HPO-2	00:00:47

edunet foundation

- **Ridge:**

  -A linear regression model that adds a regularization term to the cost function to prevent overfitting.

  -it is worked by minimizing the sum of squared errors between predicted and actual values.

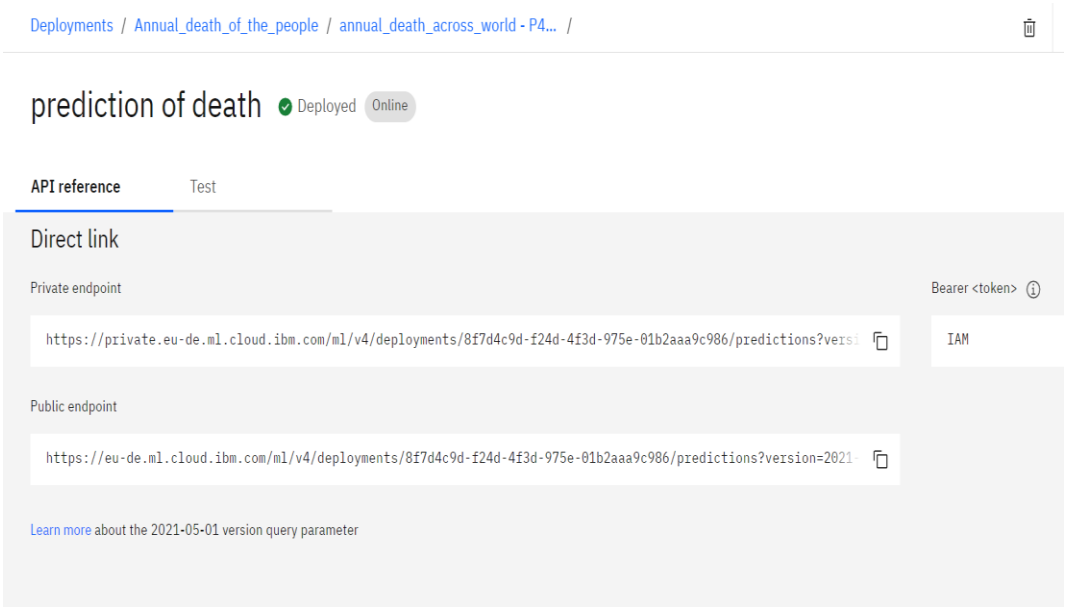
  -Adding a penalty term proportional to the square of the model's weights.-It is used when the model simplicity and interpretability and when dealing with complex, non-linear relationships between features and target variables.

  ADVANTAGES:

  --Reduces overfitting ,improves model generalization, and handles multicollinearity.

- **Deployment section:**

  Deloyment refers to the process of making machine learning models,AI workflows or data processing pipelines availabe for production.



- **Data Input:**

  Input features used by the algorithm such as annual death by ages ,different countries(entity),code,year(1950-2023), Deaths - Sex: all - Age: Variant: estimates

- **Training Process:**

  cross-validation contains Root mean squared error(RMSE),Mean Absolute Error(MAE),Mean Squared Error(MSE),Explained (variance),Median Absolute Error (MedAE) etc. Mostly RMSE is optimized in algorithm analysis for pipeline comparison.

- **Prediction Process:**

  The predicted columns in a analysis of annual death by age model predicts Multiclass classification and Regression type.

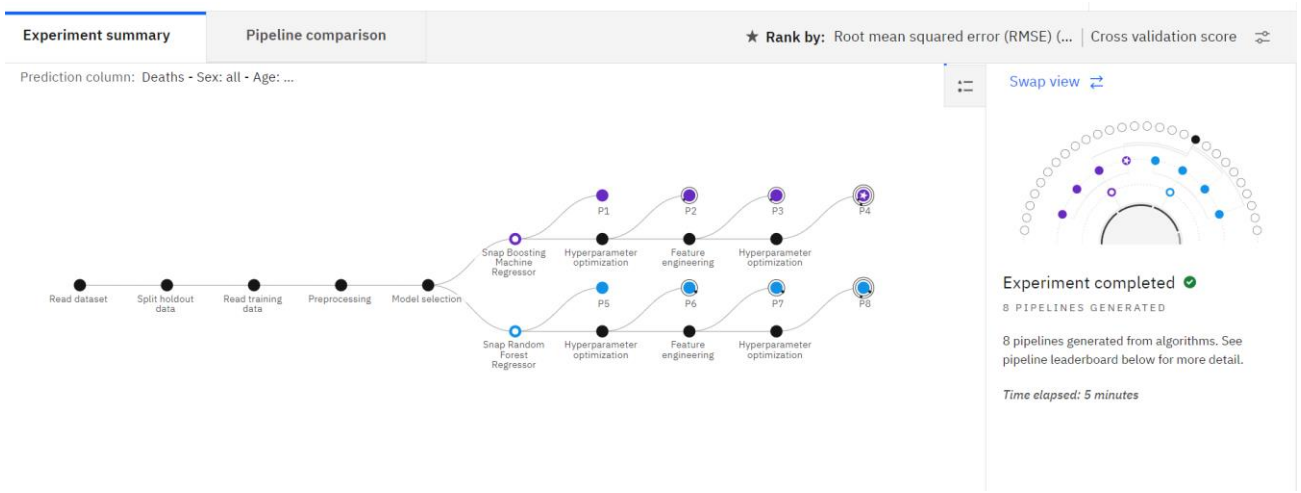
  Multiclass Classification: Output variable is categorical with more than two classes.The goal is to predict the class label for a new instance.

  eg: Entity, code from the analysis.

  Regression :The goal is to predict multiple continuous values for a new instance.

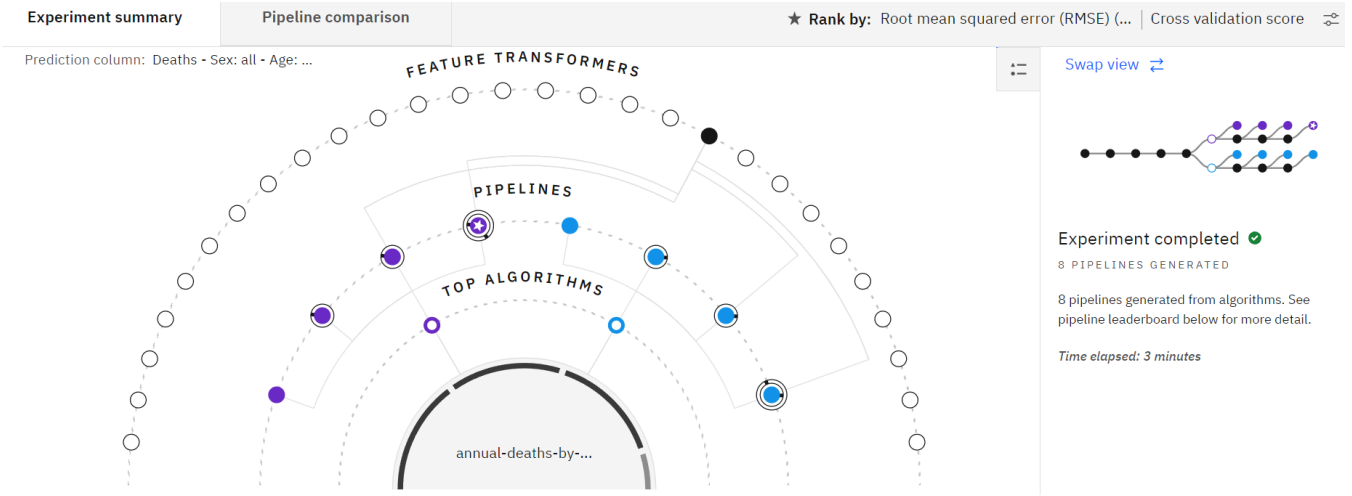
  eg: year(int) Deaths - Sex: all - Age: 0-4 - Variant: estimates(dec) etc..

RESULT



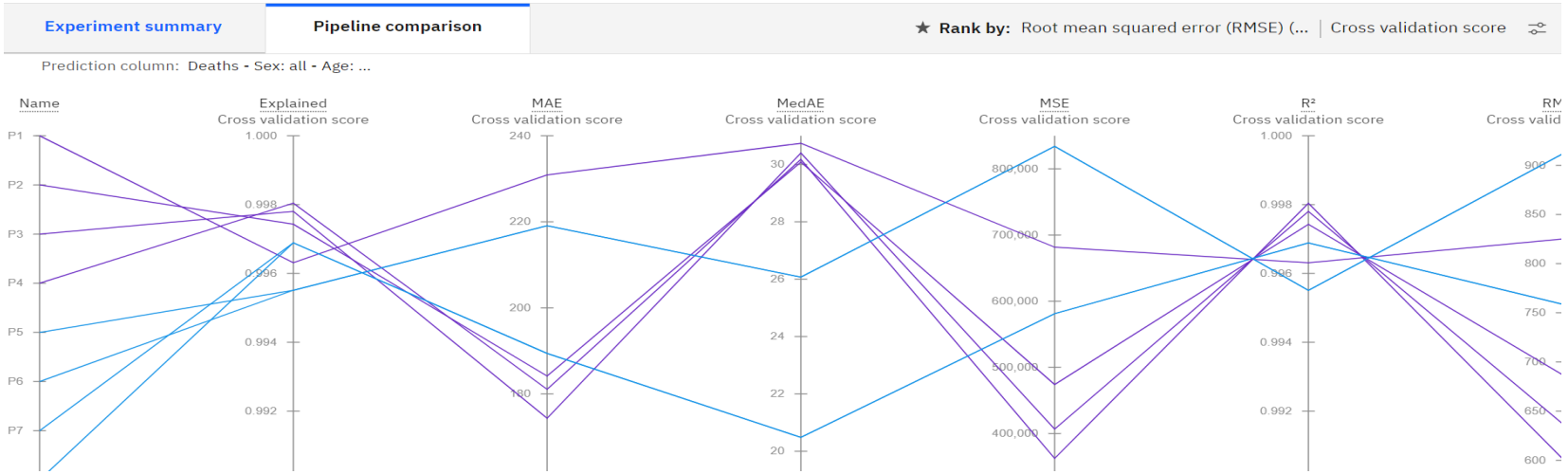
- The picture represents the progress map in experiment summary.
- 8 pipelines are generated from algorithm.

- The following represents the Relationship map.
- It contains different states like feature transformers,pipelines,top algorithms,



edunet foundation

- The following images shows the pipeline comparison of algorithm.



Experiment summary | Pipeline comparison | ★ Rank by: Root mean squared error (RMSE) (... | Cross validation score

	Rank	Name	Algorithm	RMSE (Optimized) Cross Validation	Enhancements	Build time
★	1	**Pipeline 4**	Snap Boosting Machine Regressor	**601.957**	HPO-1 FE HPO-2	00:01:07
	2	**Pipeline 3**	Snap Boosting Machine Regressor	**637.504**	HPO-1 FE	00:00:53
	3	**Pipeline 2**	Snap Boosting Machine Regressor	**687.081**	HPO-1	00:00:08
	4	**Pipeline 8**	Snap Random Forest Regressor	**758.833**	HPO-1 FE HPO-2	00:01:12
	5	**Pipeline 7**	Snap Random Forest Regressor	**758.833**	HPO-1 FE	00:00:40
	6	**Pipeline 1**	Snap Boosting Machine Regressor	**824.944**	*None*	00:00:01

CONCLUSION

- This project leveraged machine learning and auto AI to analyze and predict annual death rates by age,By utilizing a range of algorithms and techniques ,including Random Forest Regressor and Ridge Regressor,we achieved hidh accuracy in predicting death rates across different groups.
- Identification of high risk age groups and associated factors, informing public health initiatives.

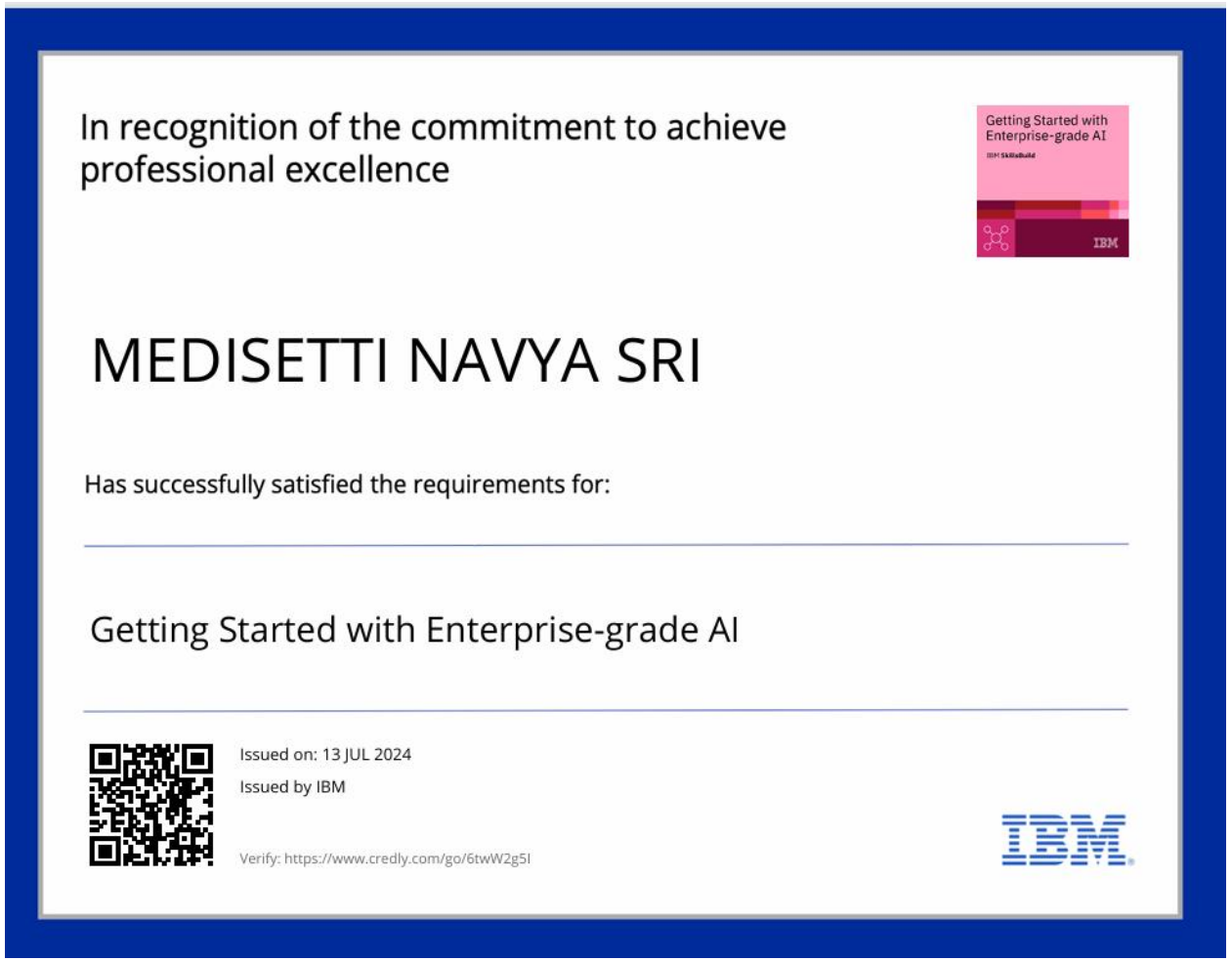
FUTURE SCOPE

- optimizing the algorithm for better performance and easy to predict analytics, population health management, Healthcare resource organizations, Mortality rate benchmarking. Identifying individuals with high mortality risks to provide early palliative care interventions.
- Using machine learning algorithm to forecast future death rates based on historical trends and demographic changes.
- By using machine learning it is easy to enhance the data visualization and communication.

REFERENCES

- https://ourworldindata.org/grapher/annual-deaths-by-age?tab=table&time=latest
- https://cloud.ibm.com/services/data-science-experience/crn%3Av1%3Abluemix%3Apublic%3Adata-science-experience%3Aeu-de%2F4419b6d5863247139105055 0cc9dd8b1%3A586b783e-5632-44f7-8c70-edcb888f66b3%3A%3A?paneId=manage
- https://eu-de.dataplatform.cloud.ibm.com/home2?context=cpdaas&apps=data_science_experience&nocache=true&onboarding=true&quick_start_target=data_science_experience
- Meta AI, chatgpt.com and google.com are used.

COURSE CERTIFICATE 1



In recognition of the commitment to achieve professional excellence

Getting Started with Enterprise-grade AI

IBM SkillsBuild

IBM

MEDISETTI NAVYA SRI

Has successfully satisfied the requirements for:

Getting Started with Enterprise-grade AI

Issued on: 13 JUL 2024

Issued by IBM

Verify: https://www.credly.com/go/6twW2g5I

IBM

edunet foundation

COURSE CERTIFICATE 2

edunet foundation

THANK YOU

edunet foundation