# Predicting individuals at the risk of becoming NEET (Not in Education, Employment or Training)

Andrew Kitchen
Project Manager

Aygul Zagidullina
Technical Mentor

Nazeefa Muzammil
Fellow

Yannick Stadtfeld
Fellow

Mahima Tendulkar Fellow **Emmanuel Owusu Ahenkan** 

Fellow











## PROBLEM STATEMENT

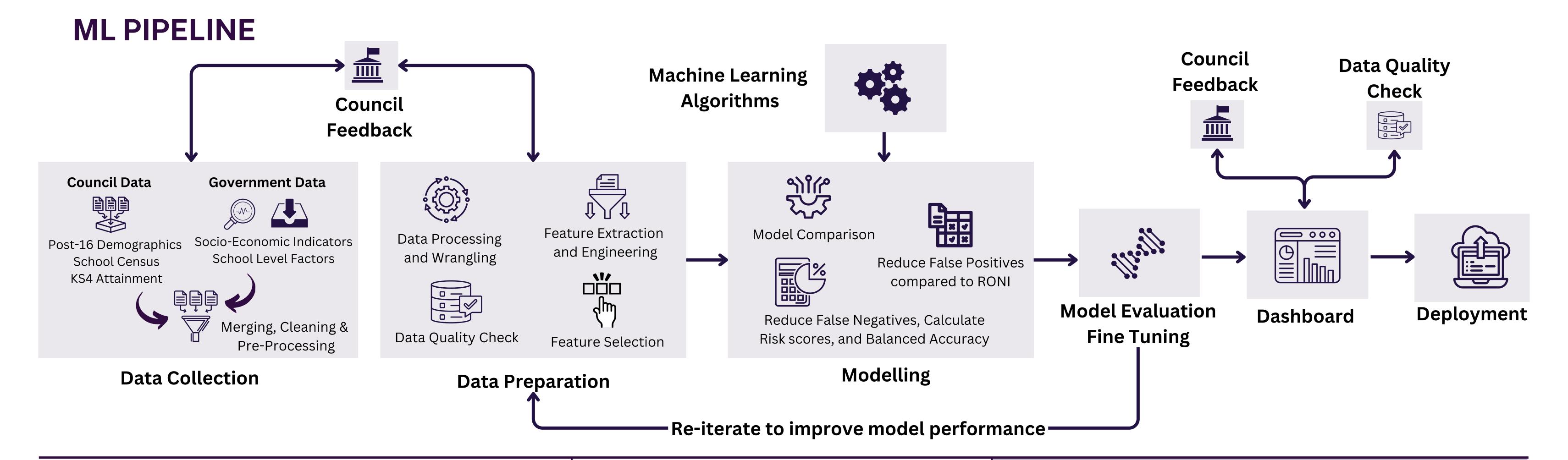
- Every year ~2-3% of young people become NEET.
- Being NEET can lead to poorer quality of life, lower income and mental health issues.
- Support of NEET individuals over their lifetime is costly to public finances.
- There is no large-scale data-driven solution for decision-making.
- Interventions by careers professionals could be more efficient with better intelligence.

# **OBJECTIVES**

- Identify datasets and build a robust processing pipeline that scales nationally.
- Develop predictive models to identify individuals at risk of becoming NEET.
- Create an open-source visualisation tool to analyse predictions at a council, school and individual level.

BASELINES	RONI	NEET'22	NEET'23
Post-16 Demographic			
GCSE Attainment	×		
School Census	X		
Regional Socio-Economic Indicators	X	X	
School Level Factors	×	×	
<b>Automated Predictions</b>	X		
Data Quality Checks	X	?	
Minimise False Negatives	×		
Feature Selection/Engineering	×	×	
Open Source Dashboard	X	×	
Scalable Deployment Solution	×	×	

## **METHODOLOGY** Year 11 Year 12 Year 13 (age 15 to 16) (age 16 to 17) (age 17 to 18) Home **About the young person** Council **Classifier Models:** Establishments Currently in Year 12 (Cohort 2019-20) Logistic Regression Young persons Risk of NEET ② RONI Score ② KS4 Attainment 8 Data & Predictions LightGBM 76% 4.7% 67.0 nstructions Mar Sept ilter by KS4 establishment UID: 246178 Establishment (KS4): Oasis Academy Lister Park Enrolment Status (KS4): C GCSE Post-16 Post-16 Post-16 Post-16 Gender: Female ilter by predicted risk (percent) SEN Provision: SEN Support Demographics Demographics Demographics Demographics Challenge: highly imbalanced rimary SEN need: Physical Disability Postcode: BD9 4Y3 Secondary SEN need: None data (2% NEET vs 98% non-NEET) John Doe (246178) Post-16 activities (NCCIS) **Solution:** use SMOTE (-NC) for Model 1 Model 3 Model 2 Current activity: Full-time education - further education training and hyper-parameter tuning



# **IMPACT**

- No individual is overlooked due to the nondiscriminatory methodology.
- Councils can target planning and delivery of interventions more efficiently in collaboration with schools.
- Scaling the tool at a national level could lower the percentage of NEET individuals within England, providing significant societal benefits (School Year 12-13).

# RESULTS

- A prototype interoperable pipeline for processing school data for modelling purposes.
- New models outperform both baseline models in terms of Accuracy and False Negative rate.
- The dashboard provides analysis of risk factors on an individual, school and council level.

# **FUTURE WORK**

- Testing, adjusting and refining the tool with other council's data.
- Engineering additional features and new data augmentation to further improve performance.
- Further develop the approach to ethical considerations, generalisation and bias.
- Explore the use of centralised school data sources for efficiencies in the delivery paradigm with regard to upscaling the tool nationally.

