ADSP

CORN YIELD FORECAST

Parsa Taati

Hamed Goldoust

Afsoon Abbasi

Politecnico di Torino

FONDAZIONE links

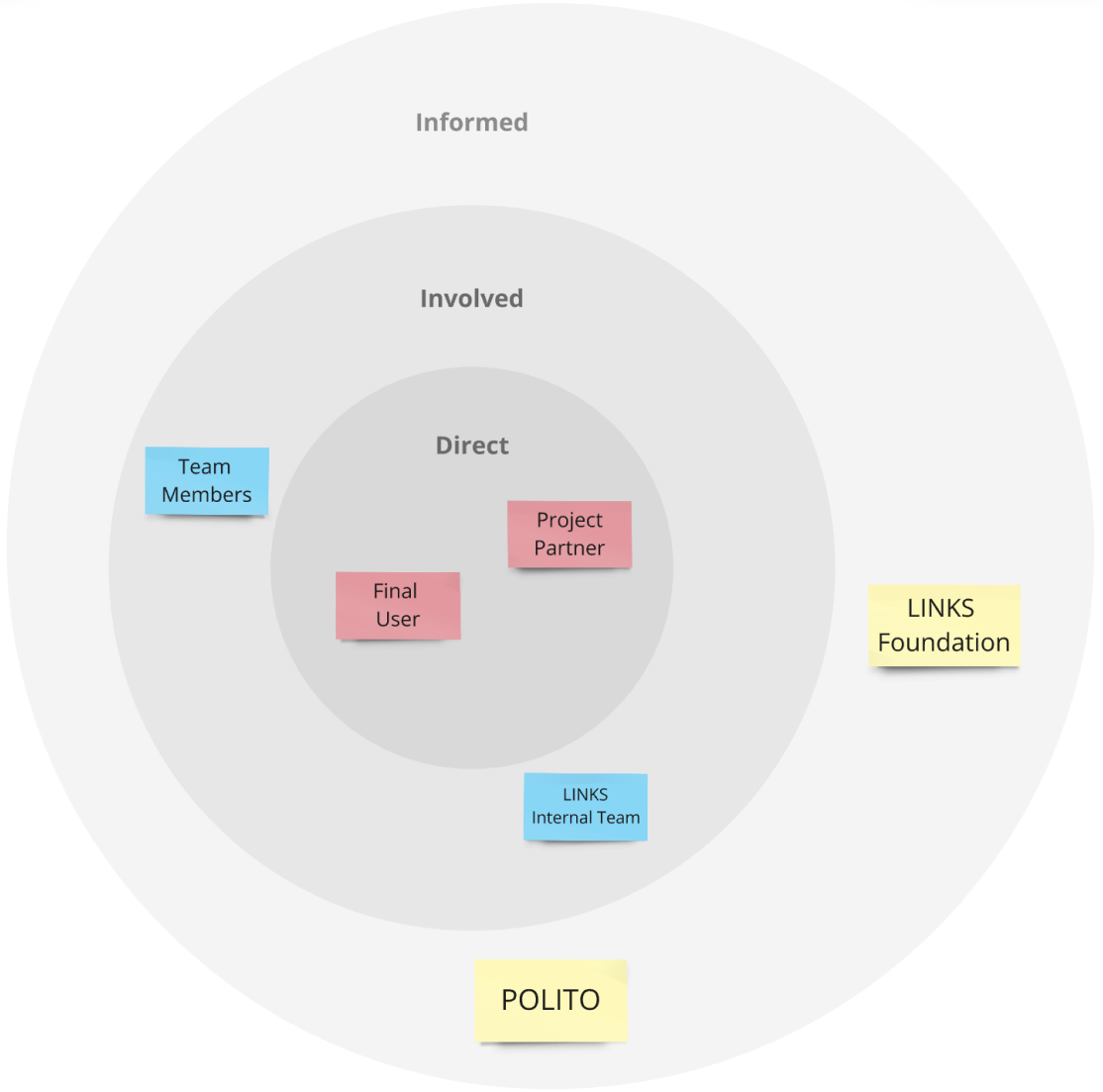
PASSION FOR INNOVATION

Project Background

- Purpose of the Project:
  - Introduction to crop yield forecasting within the European Guardians project, aiming for sustainable agricultural practices.

- Scope and Locations:
  - Piedmont, Italy
  - Spain, focusing on the data from the Spanish pilot site.

- Stakeholders map



Name: Giuseppe De Luca
Age: 60
Role: Farmer (Final User)
Background: Farmer, based in Piedmont region

Behaviors & Habits
- Checks weather and inspects crops daily. Relies on experience but is open to simple, practical technology.
- Uses basic smartphone functions, prefers straightforward tools without jargon.

Pain-Points & Frustrations
- Struggles with sudden weather changes affecting crop planning.
- Frustrated by tools that are hard to use or understand.

Needs & Goals
- Wants clear advice on planting, harvesting, and pest management.
- Needs accurate weather and crop condition forecasts to optimize yield.

Data Can Produce/Use
- Provides insights into crop performance and local weather impacts.

Expected Outputs
- A mobile tool with clear, practical recommendations.
- Notifications for weather risks and crop management advice, helping him make timely decisions.

Name: Sara Ferrari
Age: 29
Role: LINKS Internal team member
Background: Data Scientist specializing in machine learning at LINKS Foundation

Behaviors & Habits
- Analytical and curious, often conducts deep dives into data to find patterns.
- Values collaboration and actively seeks feedback from team members.
- Enjoys working with cutting-edge technology and exploring its applications in real-world scenarios.

Pain-Points & Frustrations
- Frustrated by inconsistent data quality or access issues.
- Finds it challenging to align her expertise with the broader project scope without clear communication.

Needs & Goals
- Needs access to reliable data sources and clear project guidelines.
- Goal: Contribute technical expertise to develop a robust, high-performing model that meets academic and industry standards.

Data Can Produce/Use
- Analyzes and cleans data, generating processed datasets for modeling.
- Creates reports and visualizations that highlight model performance and insights derived from data.

Expected Outputs
- A functional, validated model that is ready for integration into real-world applications.
- Technical documentation on the model's design, limitations, and areas for potential improvement.

Name: Dr. Elena Rossi
Age: 55
Role: Professor and Academic advisor
Background: Professor of Data Science at Politecnico di Torino with extensive experience in AI and machine learning research.

Behaviors & Habits
- Meticulous and critical, reviews student work with a focus on rigor and academic integrity.
- Regularly seeks to publish research and encourages students to do the same.
- Enjoys mentoring students and discussing theoretical aspects but insists on practical, well-documented results.

Pain-Points & Frustrations
- Frustrated by poorly organized documentation and lack of understanding of theoretical underpinnings in student projects.
- Struggles with balancing time between teaching, research, and advising student projects.

Needs & Goals
- Needs clear progress updates and well-documented work from students to assess their understanding.
- Goal: Ensure that the project reaches a level of quality suitable for publication or future research opportunities.

Data Can Produce/Use
- Provides academic resources, including research articles, past project examples, and theoretical frameworks.
- Shares insights from her research to guide students in model selection and methodological choices.

Expected Outputs
- A final project report that is academically rigorous and potentially publishable.
- A working prototype or model with detailed explanations of methodologies used, limitations, and potential future work.

Objectives and Goals

- Goal
  - Develop a model to forecast corn yields, helping farmers plan and optimize crop management.

- Broader Impact
  - Contribution to resource efficiency, climate neutrality, and emission reduction in agriculture.

General Strategy

Research & Exploration: Study maize yield factors and review relevant models.

Data Preparation: Clean, normalize, and split data.

Feature Extraction: Identify key variables impacting yield.

Model Development: Build a time series prediction model.

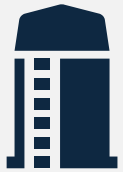
Evaluation: Measure accuracy with metrics and visualize results.

Optimization: Fine-tune model with hyperparameter adjustments.

Feedback & Improvement: Refine model based on feedback and results.

Data Overview

Data Collection: A summary of data from Spain, covering 12 years of maize crop data.

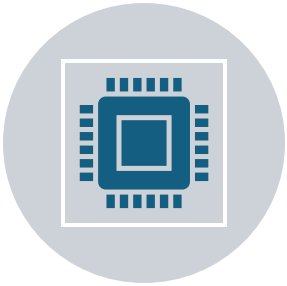


Variables Included: Yield (fresh and dried), crop variety, weather conditions, and growth cycle stages.

Methodology

- **Data Analysis Approach**
  - *Objective*: Analyze how weather conditions throughout the growing season impact the corn yield.
  - *Method*: Use time-series analysis to explore patterns and correlations between specific weather variables (e.g., temperature, humidity, precipitation) and crop yield.
  - *Expected Outcome*: Identify key weather factors and seasonal patterns that are most predictive of high or low corn yields, allowing for better model feature selection.

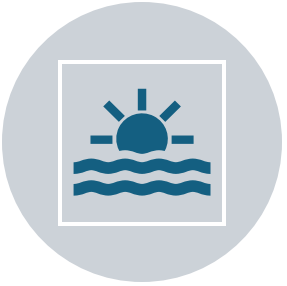
Proposed Model



Model Type: Machine learning model, specifically models that are great for time series

Software Tools: Scikit-learn will be used to build and train the Random Forest model, while data preprocessing and exploration will involve libraries like Pandas and NumPy.

Proposed Model: Introduction to a machine learning model using a Random Forest algorithm for time-series forecasting.

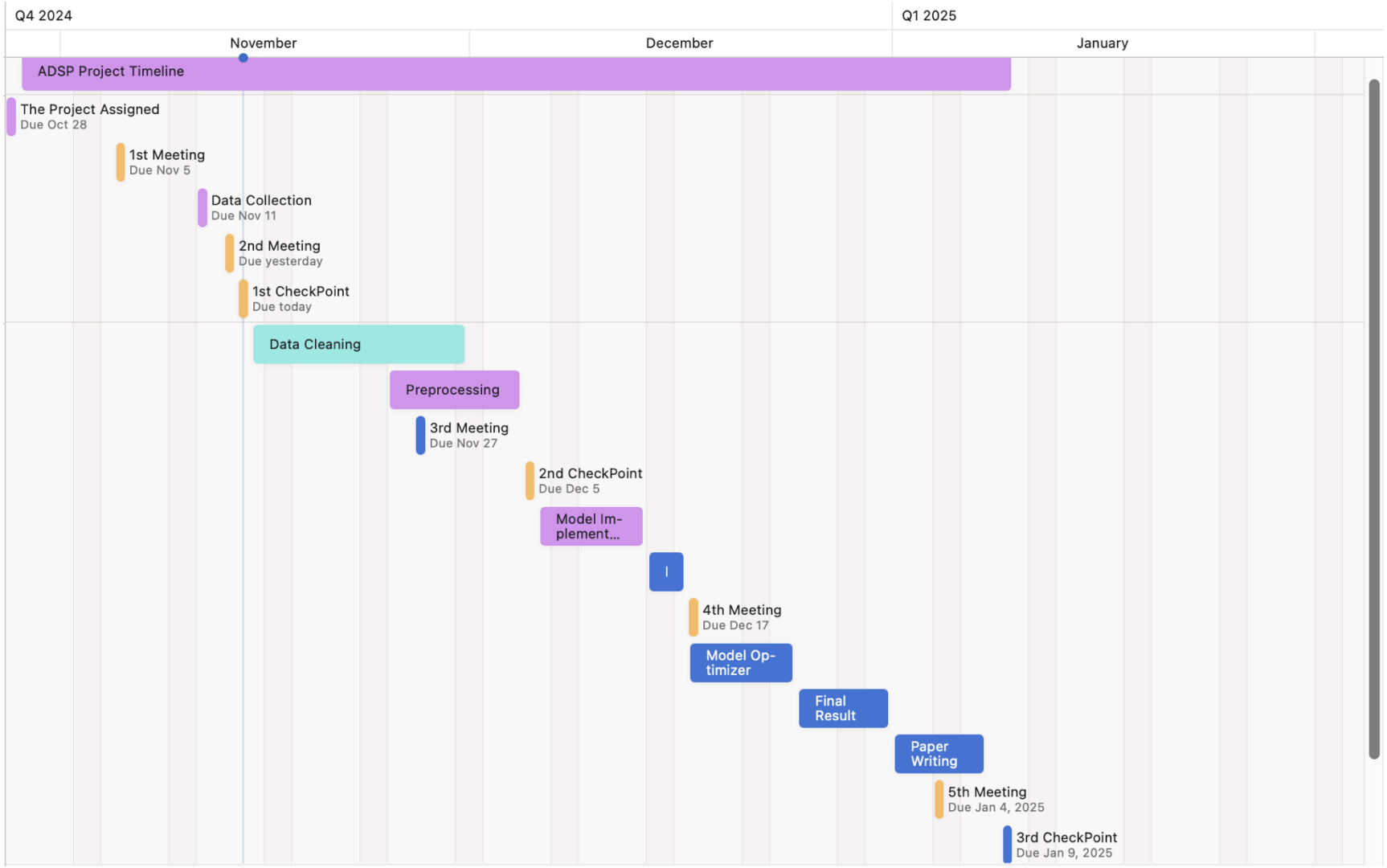


Feature Selection: Explanation of relevant weather parameters for modeling (temperature, humidity, wind, etc.).

Expected Challenges:

- Data Quality: Importance of data cleaning, handling outliers, and normalization for machine learning readiness.
- Model Training: Learning curve for machine learning frameworks, with initial focus on Scikit-learn , Google Colab, and dataset management.

Project Timeline



Thank You for your attention