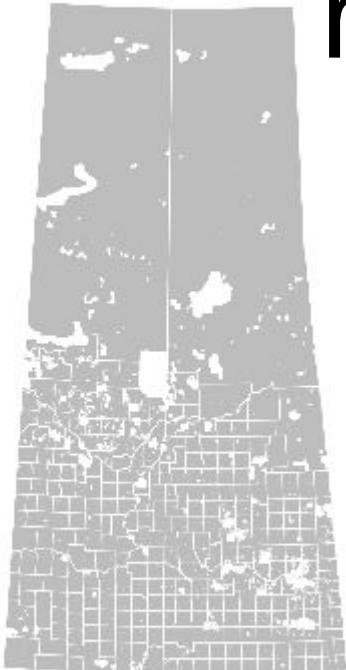


# Prediction of crop production in rural municipalities of Saskatchewan



Using Clustering and LSTM methods

*Presented by MARUF AHMAD*

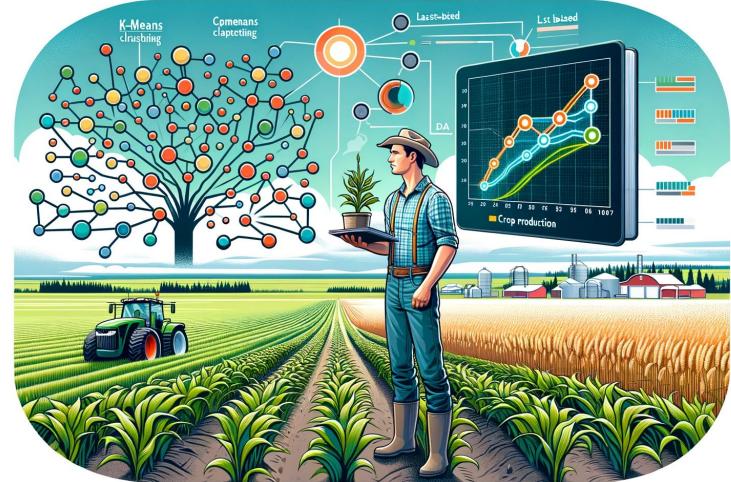
# Problem Statement

Accurately predicting crop production in rural municipalities of Saskatchewan is challenging due to the variability in climate, soil conditions, and farming practices. This uncertainty can lead to inefficient resource allocation and planning, affecting the livelihood of farmers and the agricultural economy.



# Solutions

1. Clustering: We utilized K-means clustering to identify localized crop production patterns. This helps in segmenting the data into meaningful groups for targeted analysis.
2. Prediction: For future crop production prediction, we implemented an LSTM-based Recurrent Neural Network (RNN). LSTM networks are well-suited for handling time series data and forecasting future trends.



# Crops:

Focus on two major crops



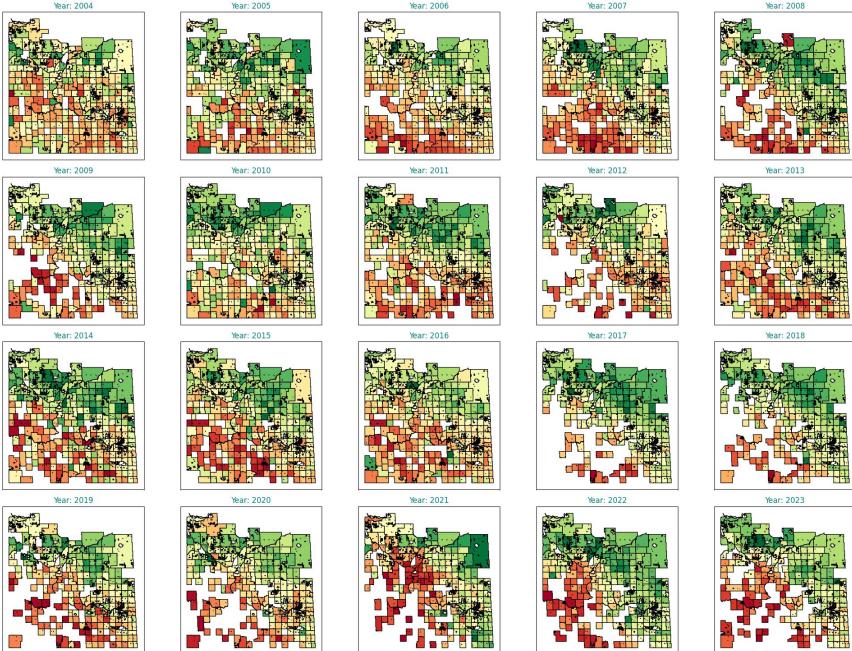
Oats



Barley

# Exploratory Data Analysis

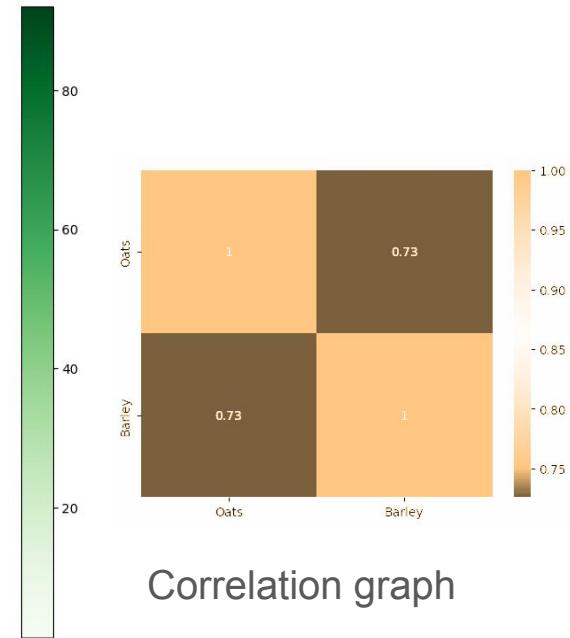
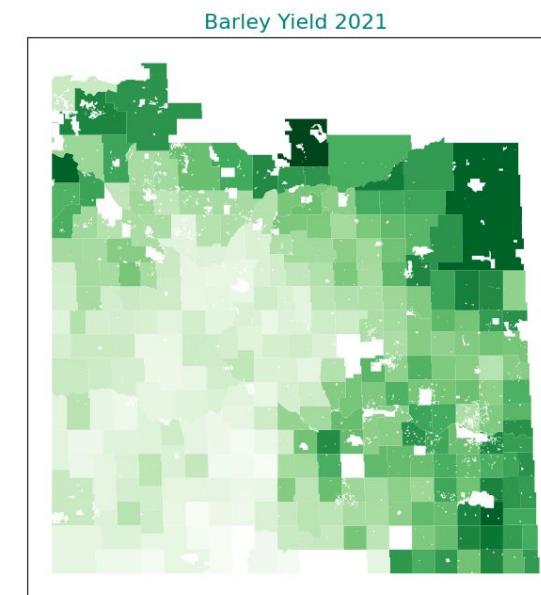
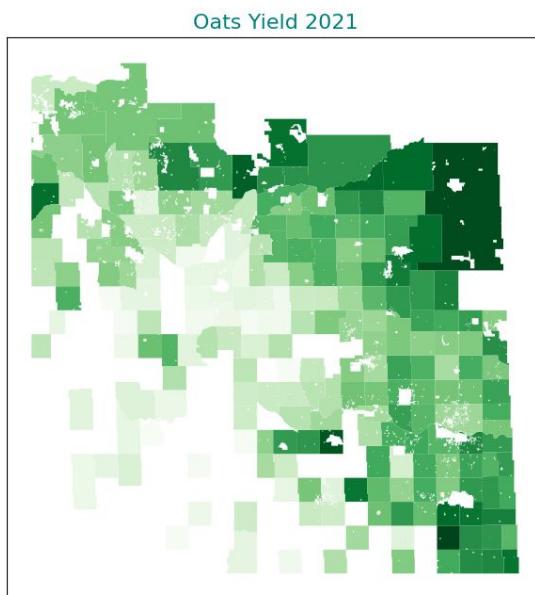
Oats Yield per Year (2004 - 2023)



Barley Yield per Year (2004 - 2023)

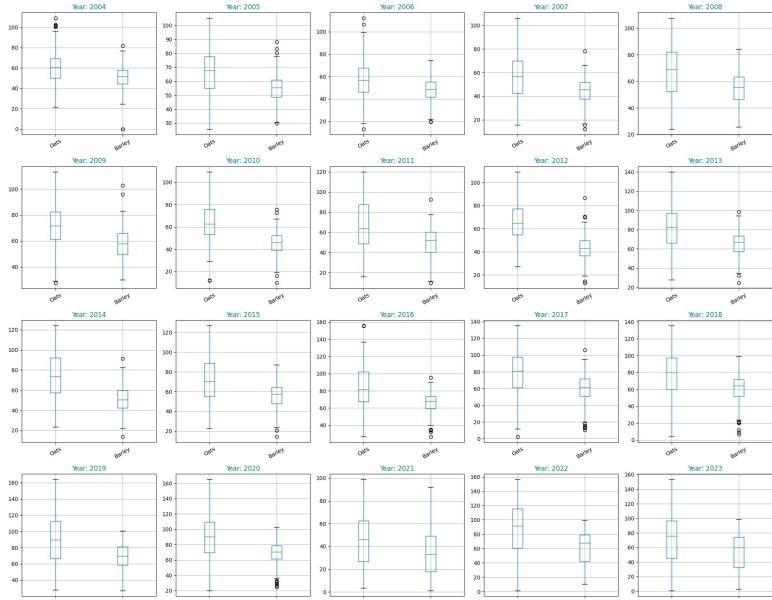


# Spatial Distribution of Oats and Barley Yields in (2021) and their correlation



- Visual representation of the correlation between different variables.
- Highlights the relationships that influence crop production.

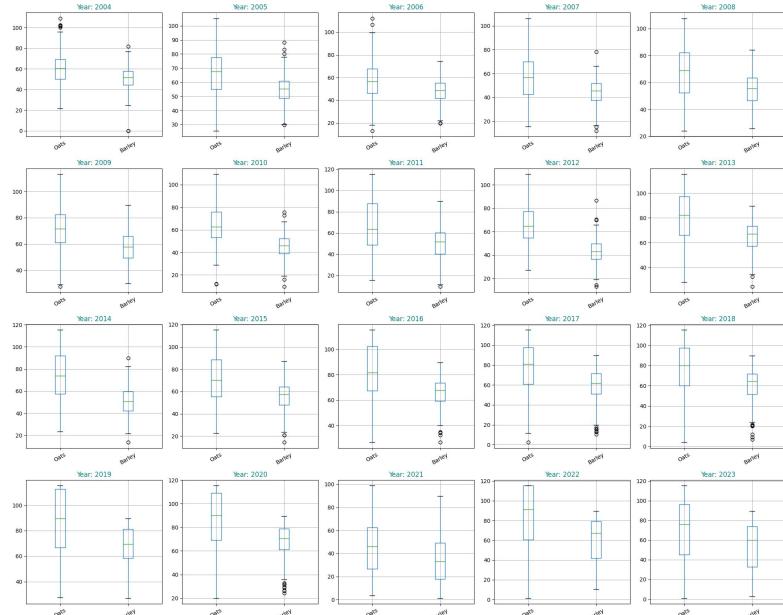
# Data Preprocessing



Before outlier Treatment

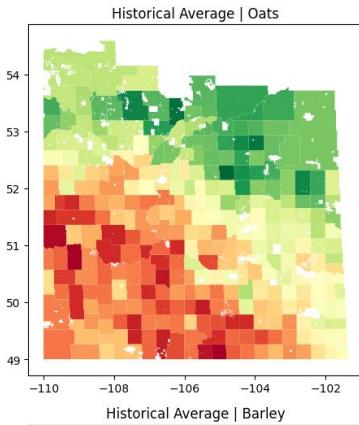
**Before Treatment:** Raw data with inconsistencies, missing values, and outliers.

**After Treatment:** Cleaned data with handle missing values, normalized values, and removed outliers for better model performance.



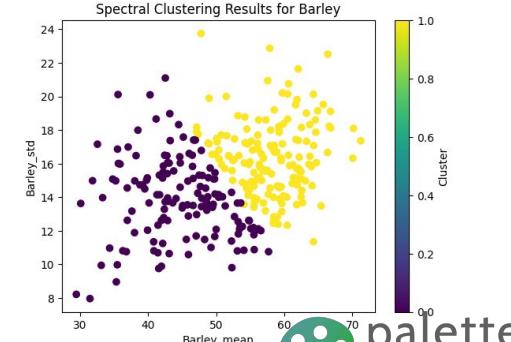
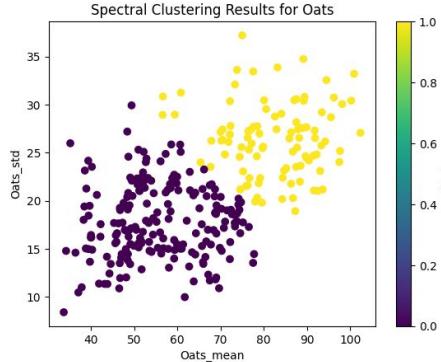
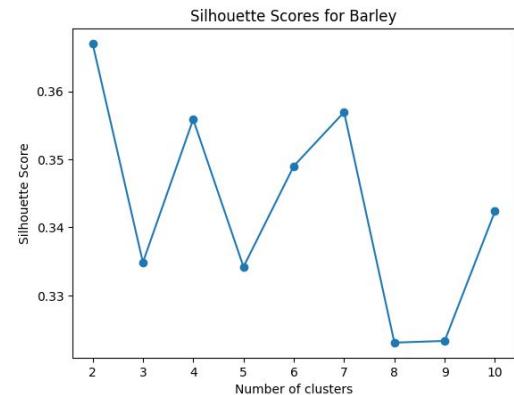
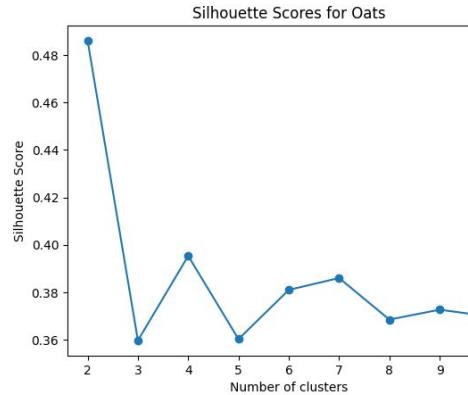
After outlier Treatment

# Cluster Analysis

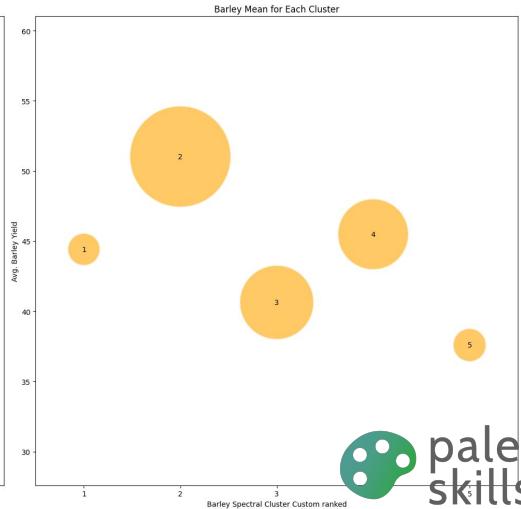
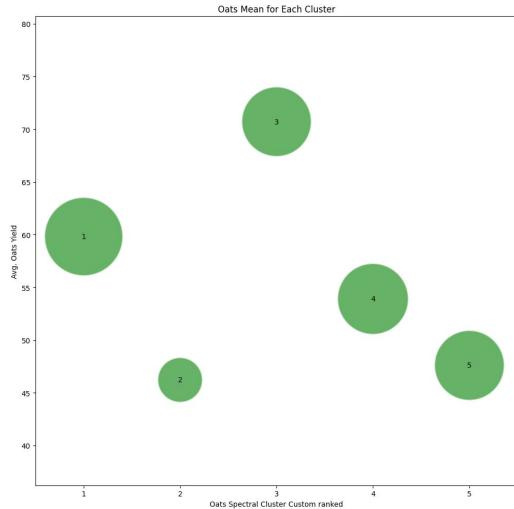
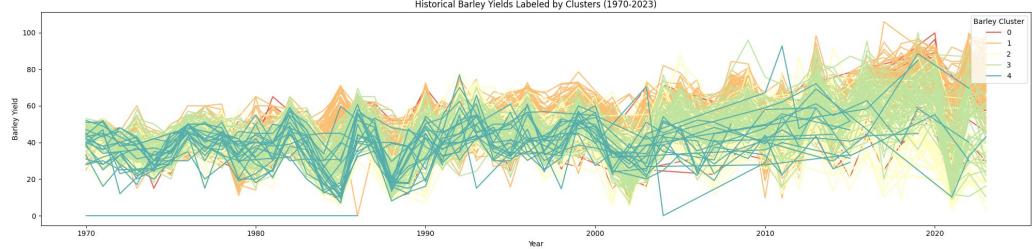
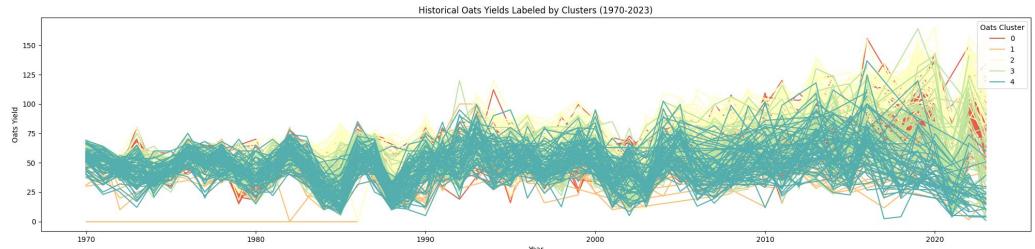
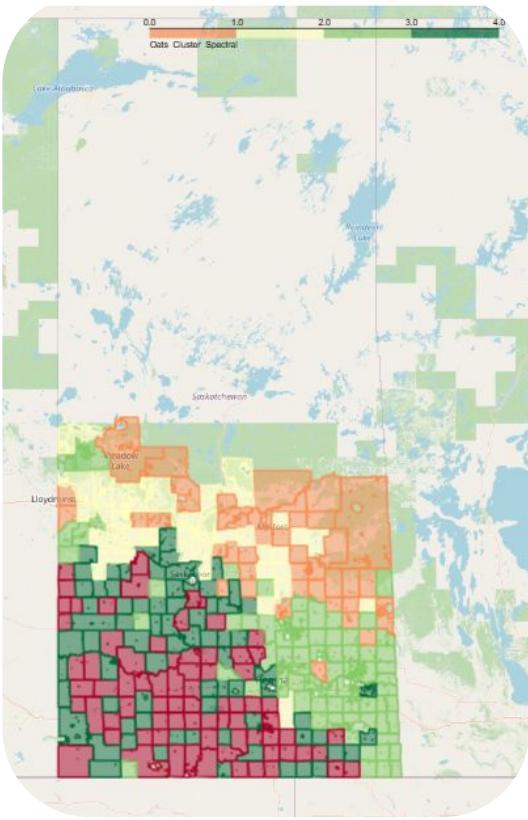


- Determined the optimal number of clusters using the Elbow method.
- Found that the best number of clusters for both oats and barley is 2.
- This clustering approach helps in understanding localized production patterns

Optimal number of clusters for both Oats and barley is 2



# Cluster Analysis

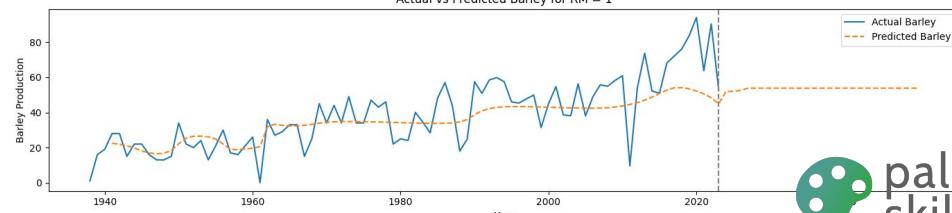
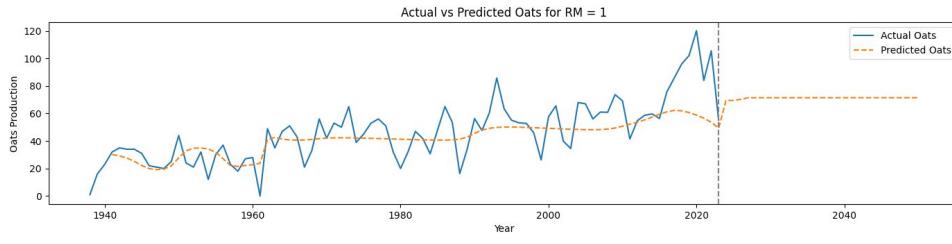
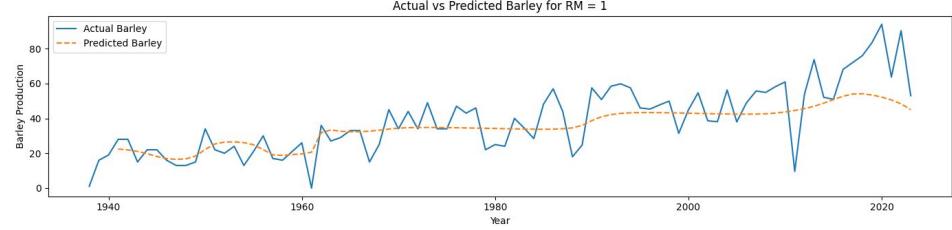
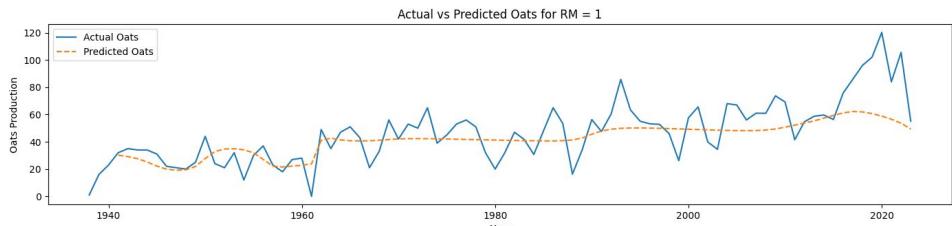
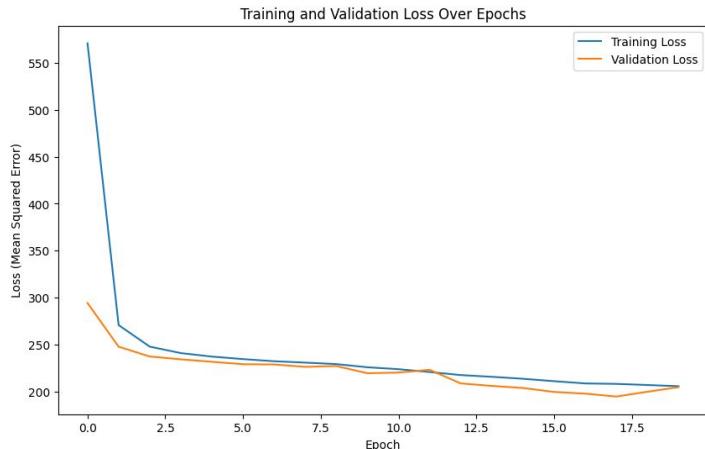


# Future Prediction

## LSTM Network Architecture:

- **Time steps:** 3
- **First layer:** 50 LSTM units with three time steps at a time
- **Second layer:** 50 LSTM units
- **Dense output layer:** 2 units
- **Optimizer:** Adam
- **Loss function:** Mean squared error
- **Epoch:** 20

LSTM is used for time series prediction, effectively capturing temporal dependencies in crop production data.



# Result & Implication

- The model achieved significant accuracy in predicting future crop production trends.
- These predictions enable better resource allocation, crop planning, and decision-making for farmers and policymakers.
- The approach supports the agricultural sector in managing uncertainty and improving productivity.



# Conclusion & future work

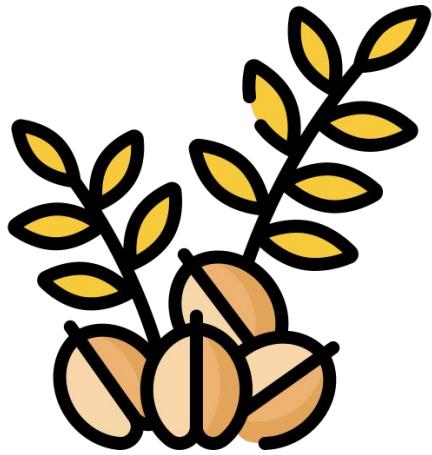
- Successfully demonstrated the effectiveness of clustering and LSTM methods in predicting crop production.
- Future work will focus on incorporating additional features such as weather forecasts and economic indicators.
- Plans to integrate real-time data for even more accurate predictions, further aiding decision-making processes.



Future is FarmGPT

# Source Code

1. Stream\_3\_Oats\_Barley\_PreprocessingNClustering  
[https://colab.research.google.com/drive/1zhNR7VjXpuh7B\\_\\_PuCIWn-TVFpt7auOe?usp=sharing](https://colab.research.google.com/drive/1zhNR7VjXpuh7B__PuCIWn-TVFpt7auOe?usp=sharing)
  
2. Stream\_3\_Oats\_Barley\_Future\_Prediction\_LSTM  
<https://colab.research.google.com/drive/1P8dH0LxyahvU4I98bbakHx952HVb0E1O?usp=sharing>



Oats

Thank You



Barley