Predict Grain Production

Final presentation - Team 11
First Semester 2023, Artificial Intelligence Project (SWE3032-41)

박혜림 (2019315581), 박정훈 (2016313844), 임동준 (2017310326),
정홍준 (2017311985), 이성준 (2017312233), 김병무 (2018313395)

Contents

Introduction (Background)

Grain is essential feature of civilization

- Modern era is not an exception!
- Even small fluctuation could affect many country's economy
    - Climate change, population, oil price, war...
- If we predict the grain production, we can act in advance!

Introduction (Objective)

Predict 4 major crops future grain production!

- corn, rice, spring wheat, winter wheat (Half of humanity eats it)
- Selected based on global production of each crops
- Most major crops in global food market

Introduction (Novelty)

Already existing concept, however..

1. Predict 4 major crops simultaneously (corn, rice, spring wheat, winter wheat)
2. Consider various kinds of features (37 kinds of features)
3. Make 6 major features prediction model -> create predicted data
4. Use predicted data to predict future grain production

3 & 4 concept will be explained in detail at later

Data (Dataset)

- Consist of U.S state data and global data
- Total 37 features (Cultivated area, population, CO2 PPM, Oil price, etc…)
- Covers 1900.01 ~ 2021.12
- Handle approximate 54k data (122 * 12 * 37)

Data (Features)

- U.S state data features

```
ds.columns
```

```
Index(['ID', 'state', 'temperature_avg', 'CDD', 'GDD', 'HDD',
       'temperature_max', 'temperature_min', 'precipitation', 'snow_depth',
       'snow_fall', 'sunlight_svm', 'sunlight_reg', 'crop',
       'productions_spring', 'harvested_spring', 'cultivated_spring',
       'productions_winter', 'harvested_winter', 'cultivated_winter'],
      dtype='object')
```

- Global data features

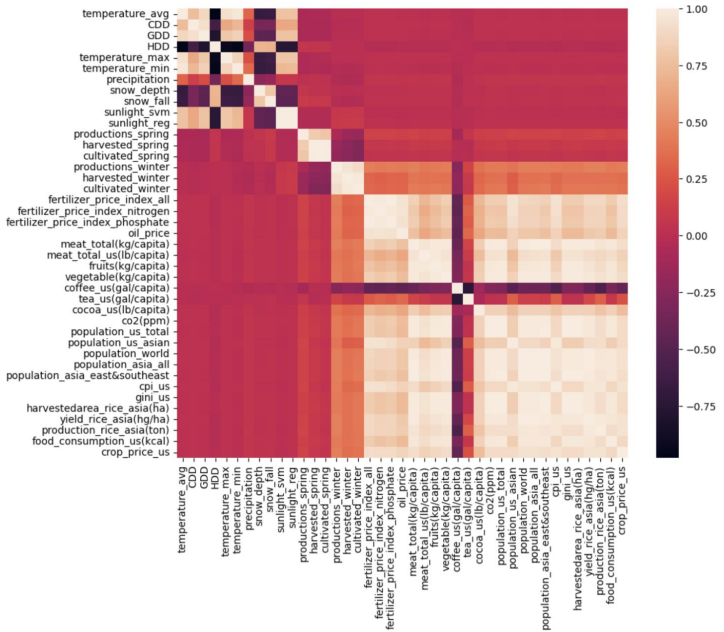
```
da.columns
```

```
Index(['ID', 'fertilizer_price_index_all', 'fertilizer_price_index_nitrogen',
       'fertilizer_price_index_phosphate', 'oil_price',
       'meat_total(kg/capita)', 'meat_total_us(lb/capita)',
       'fruits(kg/capita)', 'vegetable(kg/capita)', 'coffee_us(gal/capita)',
       'tea_us(gal/capita)', 'cocoa_us(lb/capita)', 'co2(ppm)',
       'population_us_total', 'population_us_asian', 'population_world',
       'population_asia_all', 'population_asia_east&southeast', 'cpi_us',
       'gini_us', 'harvestedarea_rice_asia(ha)', 'yield_rice_asia(hg/ha)',
       'production_rice_asia(ton)', 'food_consumption_us(kcal)',
       'crop_price_us'],
      dtype='object')
```

Data (Correlation)

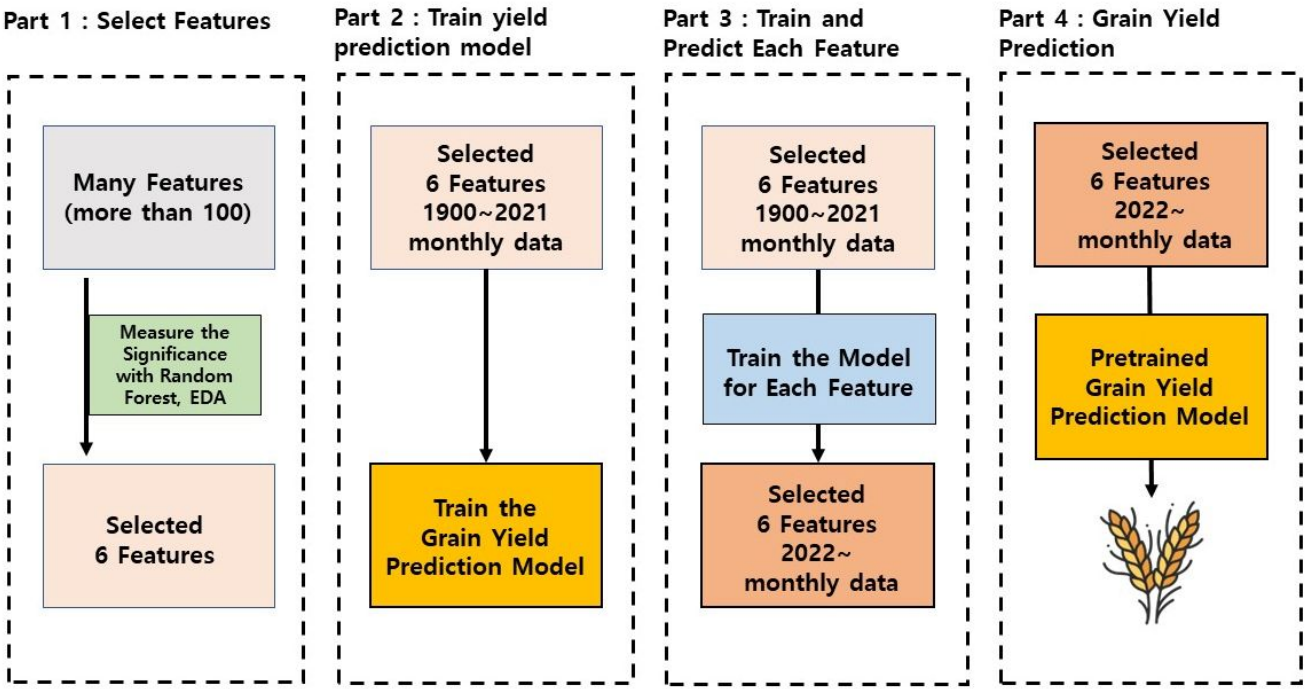
High correlation between features (bright & dark parts)

- Inspiration: maybe we can predict those features too?
- More details at later..



Model (Whole view)

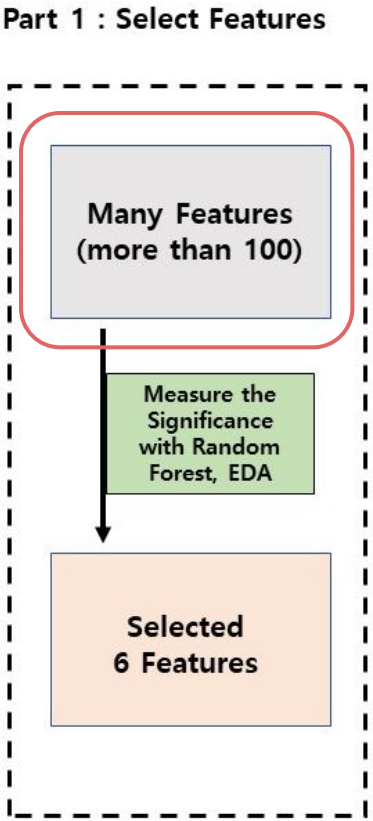
Consist of 4 parts



Part 1 (Data collection & feature selection)

Step 1. Data collection

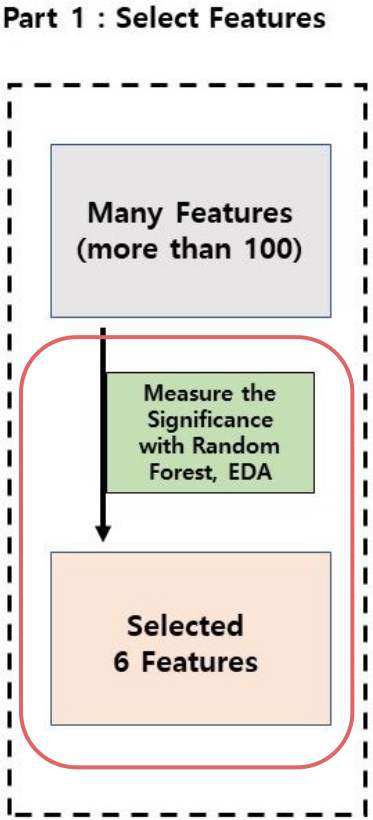
- Collect each feature's dataset
- Augment each dataset
- Fill missing values with linear interpolation



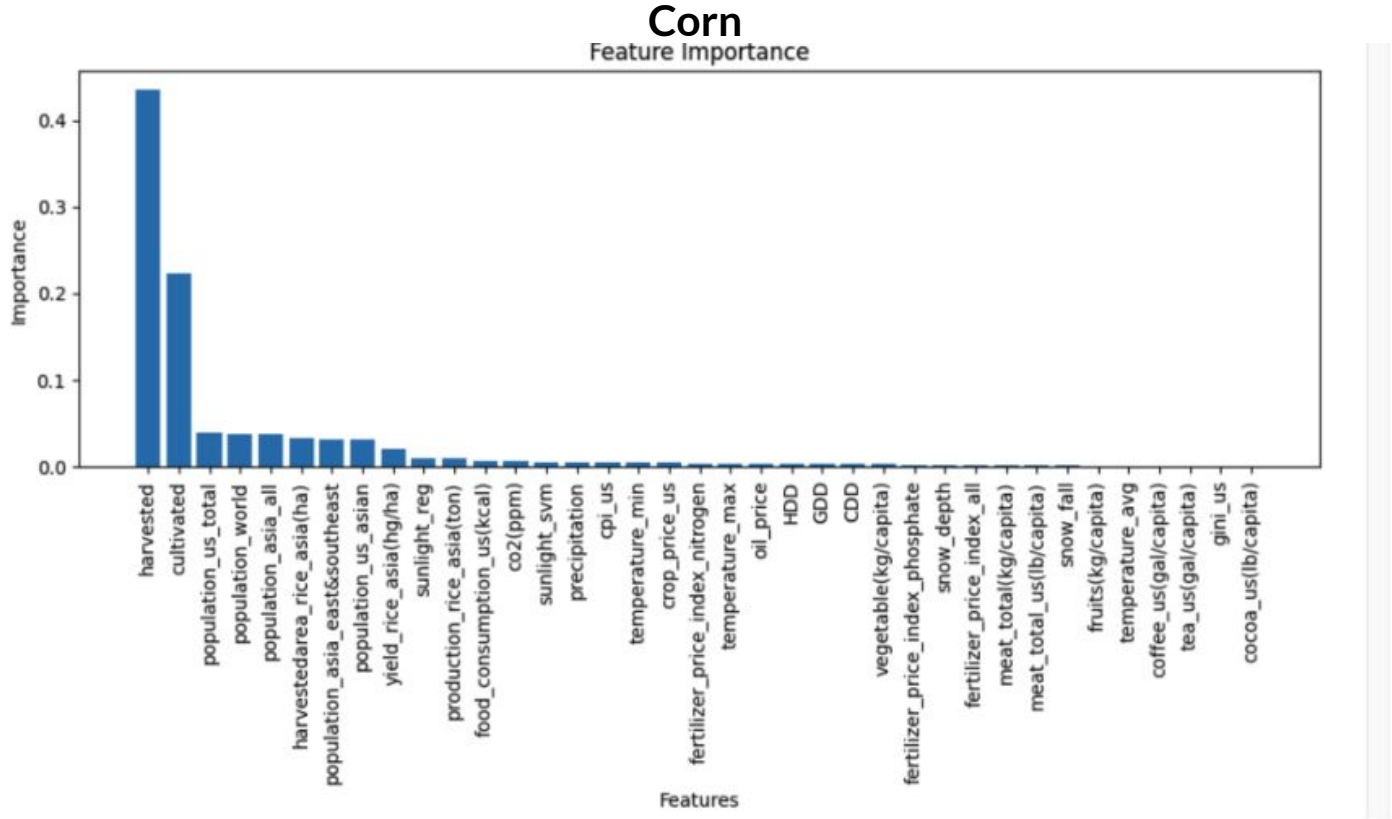
Part 1 (Data collection & feature selection)

Step 2. Feature selection

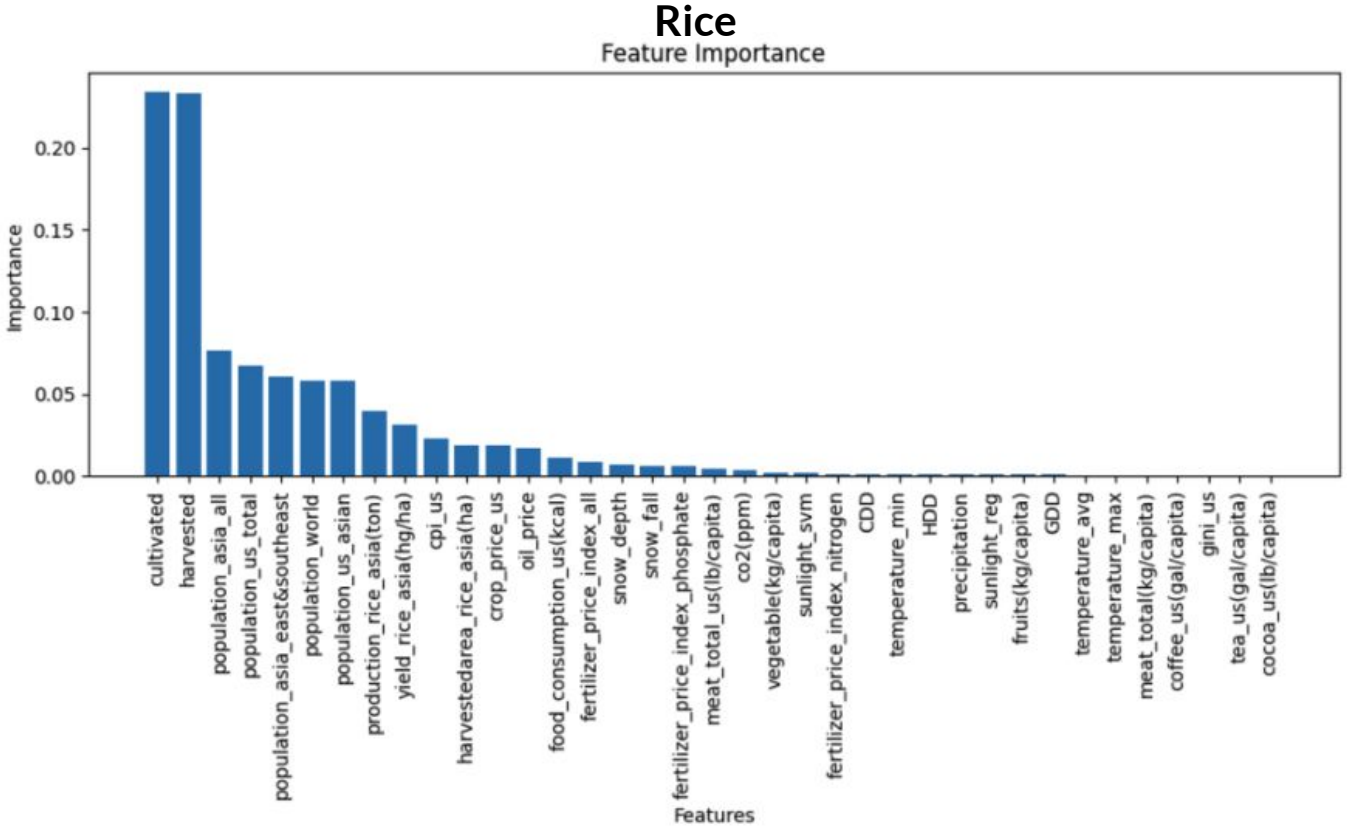
- Using random forest
  - Why? Could consider correlation between features
  - Criterion of decision tree: MSE
  - Feature importance = mean of decision trees
- Select 6 major features
- Based on feature importance



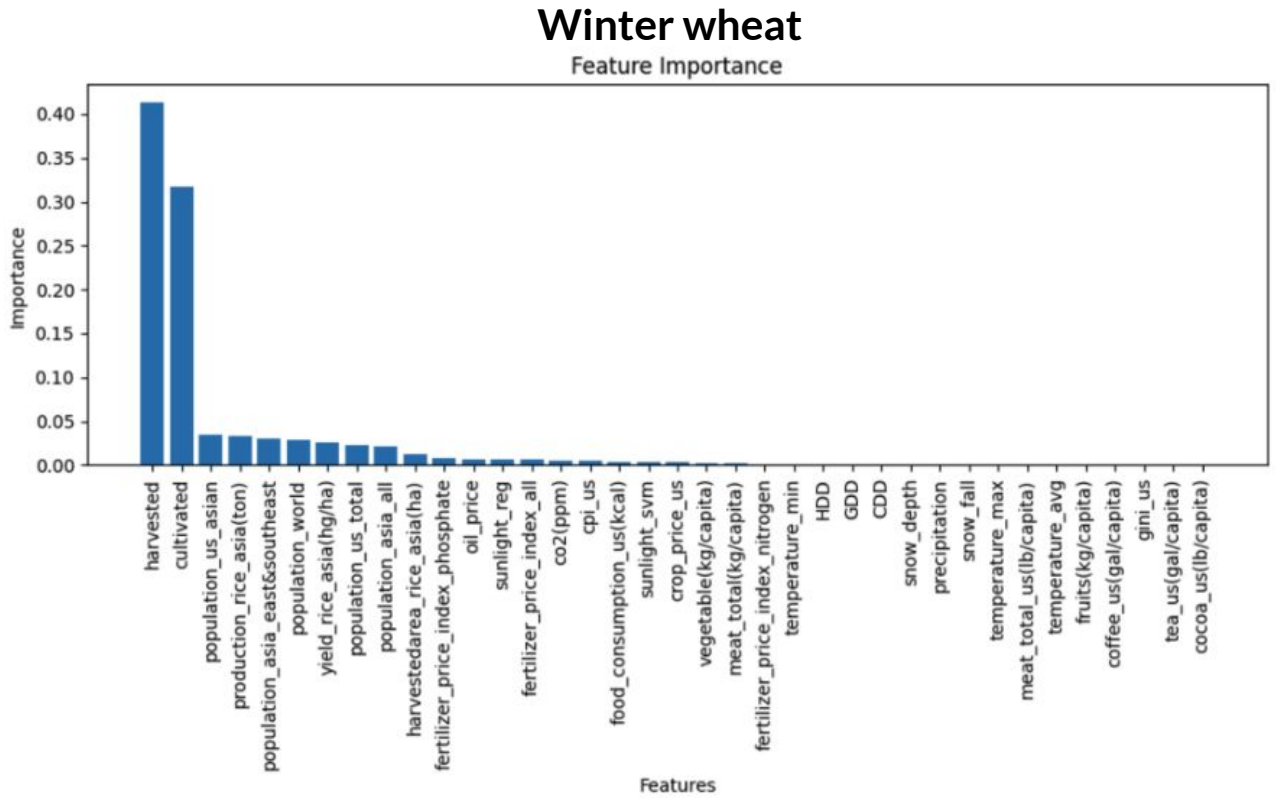
Result (Corn)


Corn
Feature Importance
Importance
0.0
0.1
0.2
0.3
0.4
harvested
cultivated
population_us_total
population_world
population_asia_all
harvestedarea_rice_asia(ha)
population_asia_east&southeast
population_us_asian
yield_rice_asia(hg/ha)
sunlight_reg
production_rice_asia(ton)
food_consumption_us(kcal)
co2(ppm)
sunlight_svm
precipitation
cpi_us
temperature_min
crop_price_us
fertilizer_price_index_nitrogen
temperature_max
oil_price
HDD
GDD
CDD
vegetable(kg/capita)
fertilizer_price_index_phosphate
snow_depth
fertilizer_price_index_all
meat_total(kg/capita)
meat_total_us(lb/capita)
snow_fall
fruits(kg/capita)
temperature_avg
coffee_us(gal/capita)
tea_us(gal/capita)
gini_us
cocoa_us(lb/capita)
Features

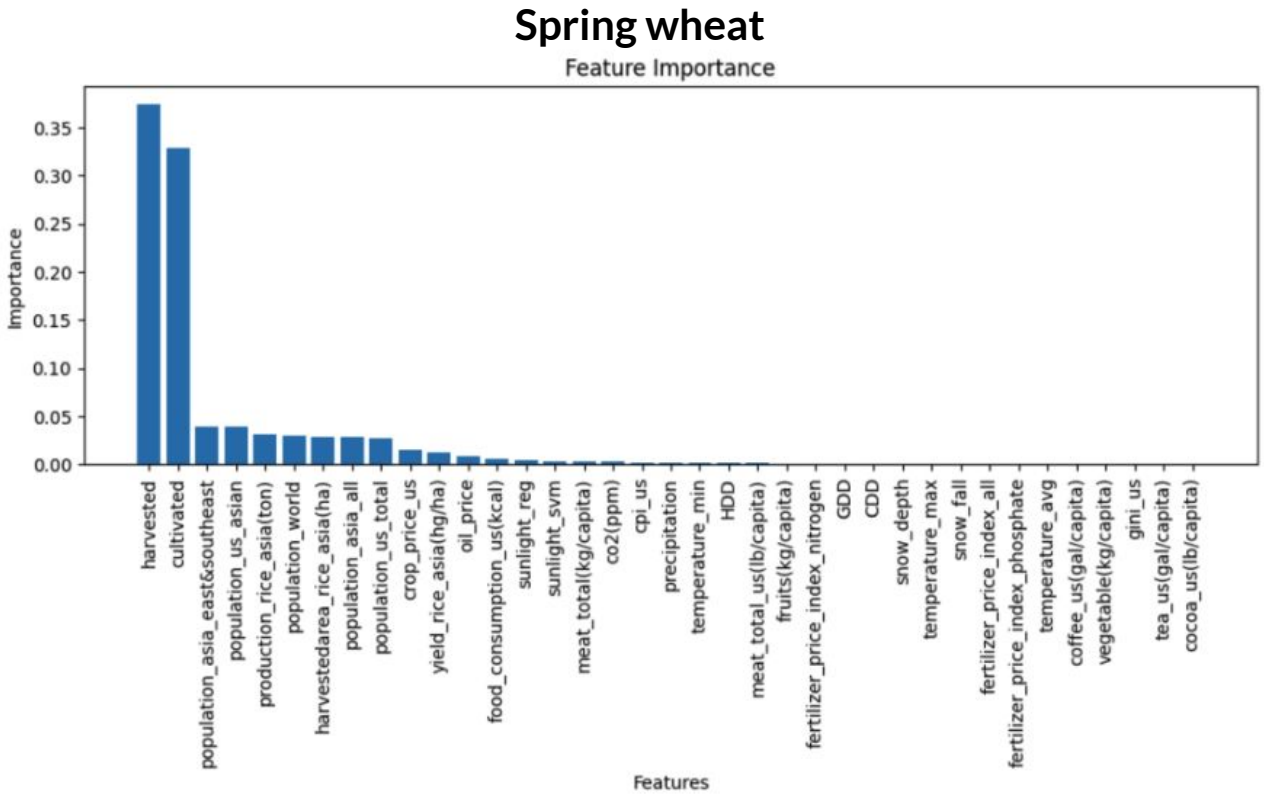
Result (Rice)


Rice
Feature Importance
Importance
0.00
0.05
0.10
0.15
0.20
cultivated
harvested
population_asia_all
population_us_total
population_asia_east&southeast
population_world
population_us_asian
production_rice_asia(ton)
yield_rice_asia(hg/ha)
cpi_us
harvestedarea_rice_asia(ha)
crop_price_us
oil_price
food_consumption_us(kcal)
fertilizer_price_index_all
snow_depth
snow_fall
fertilizer_price_index_phosphate
meat_total_us(lb/capita)
co2(ppm)
vegetable(kg/capita)
sunlight_svm
fertilizer_price_index_nitrogen
CDD
temperature_min
HDD
precipitation
sunlight_reg
fruits(kg/capita)
GDD
temperature_avg
temperature_max
meat_total(kg/capita)
coffee_us(gal/capita)
gini_us
tea_us(gal/capita)
cocoa_us(lb/capita)
Features

Result (Winter wheat)



Result (Spring wheat)



Summation

Corn

1. **Population**
2. **Sunlight**
3. **Food consumption**
4. **CO2 PPM**
5. Precipitation

Rice

1. **Population**
2. **CPI (Consumer Price Index)**
3. Grain price
4. **Oil price**
5. **Food consumption**

Winter wheat

1. **Population**
2. Fertilizer price
3. **Oil price**
4. **Sunlight**
5. **CO2 PPM**

Spring wheat

1. **Population**
2. Grain price
3. **Oil price**
4. **Food consumption**
5. **Sunlight**

Select Features

Based on frequency & novelty

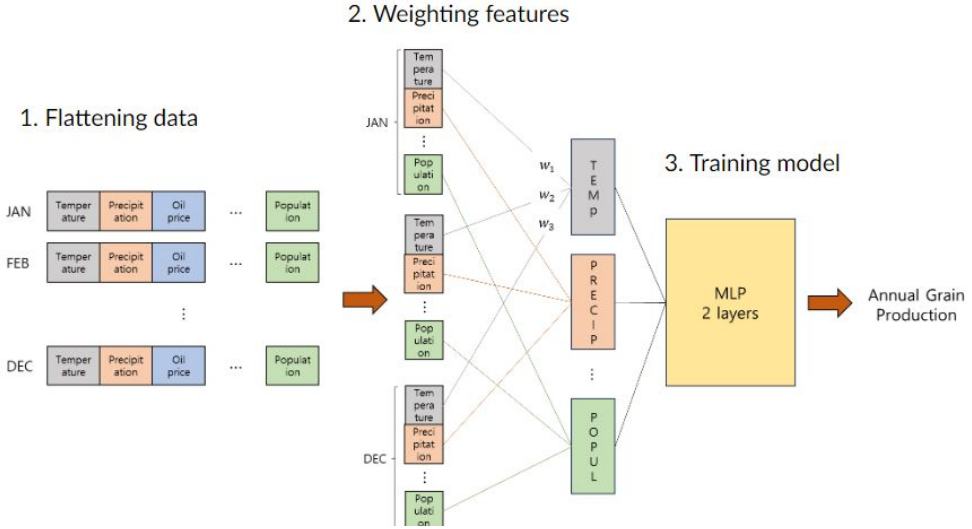
1. **Population**
2. **CPI (Consumer Price Index)**
3. **Food consumption**
4. **Sunlight**
5. **CO2 PPM**
6. **Oil price**

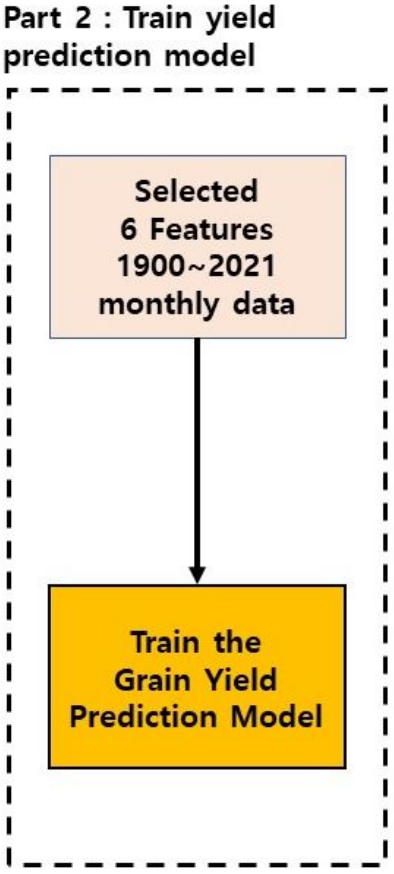
We'll predict those features at part 3!

Part 2 (Training model)

Training grain production prediction model

- Training with our dataset (all 37 features)
- Using MLP



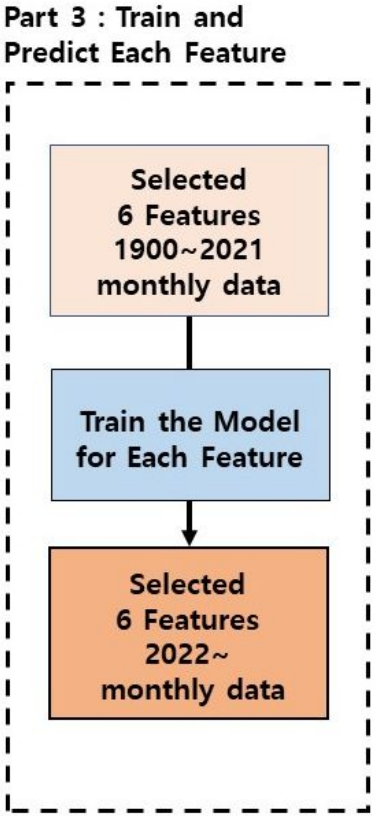


Part 3 (Get future data of 6 major features)

By training each feature prediction model!

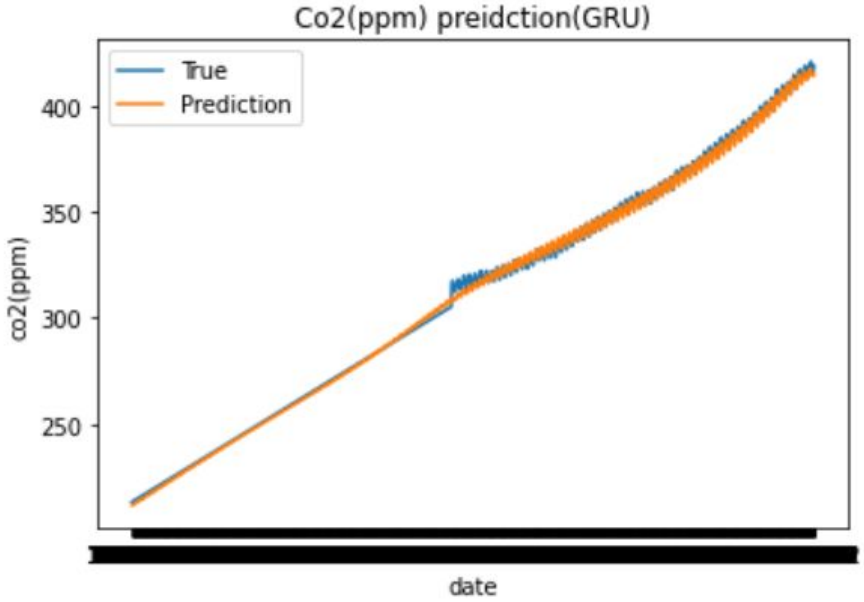
Model	Population	US CPI	Oil	Sun light	Food Consume	co2
LSTM	0.13	0.52	3145	797	3.3	2.24
GRU	-	-	3161	767	-	2.05
RNN	-	-	4717	4882	-	3.77
ARIMA	3.7	4.2	-	-	7.6	-
Ensemble(RNN,LSTM, GRU)	-	-	1531	794	-	2.20
Ensemble(LSTM, GRU)	-	-	2829	786	-	3.39

*Selected based on test loss



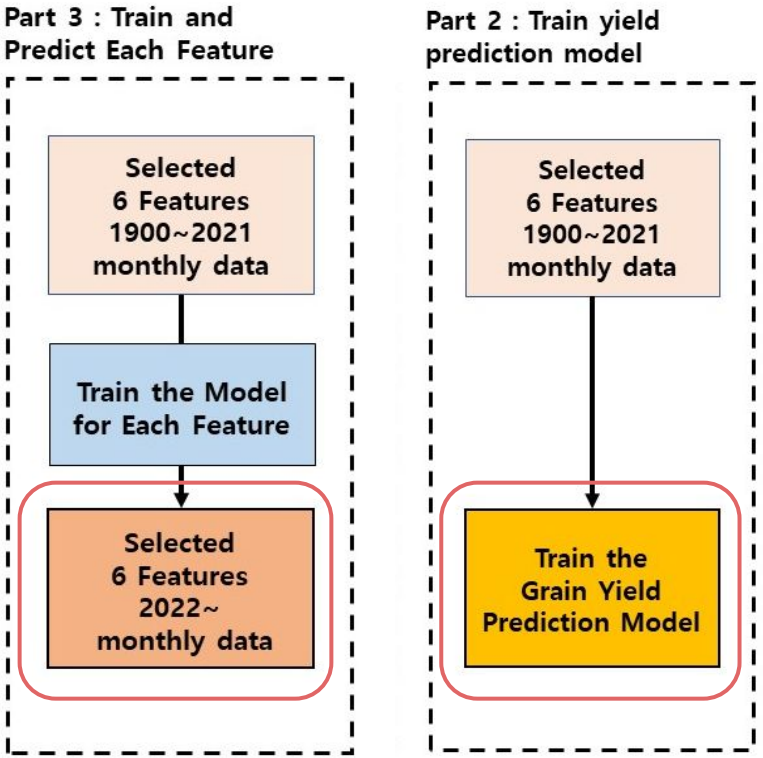
Part 3 (Get future data of 6 major features)

One of the result (co2 ppm)



What do we have now?

1. Grain production prediction model (Part 2)
2. 6 major features prediction data (Part 1, 3)

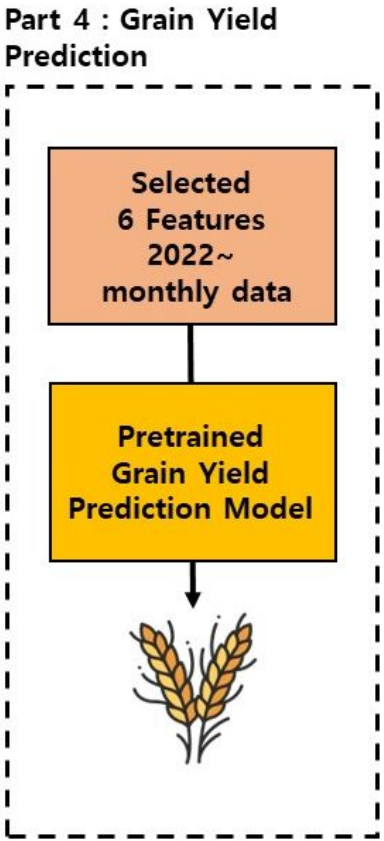


Part 4 (Predict future grain production!)

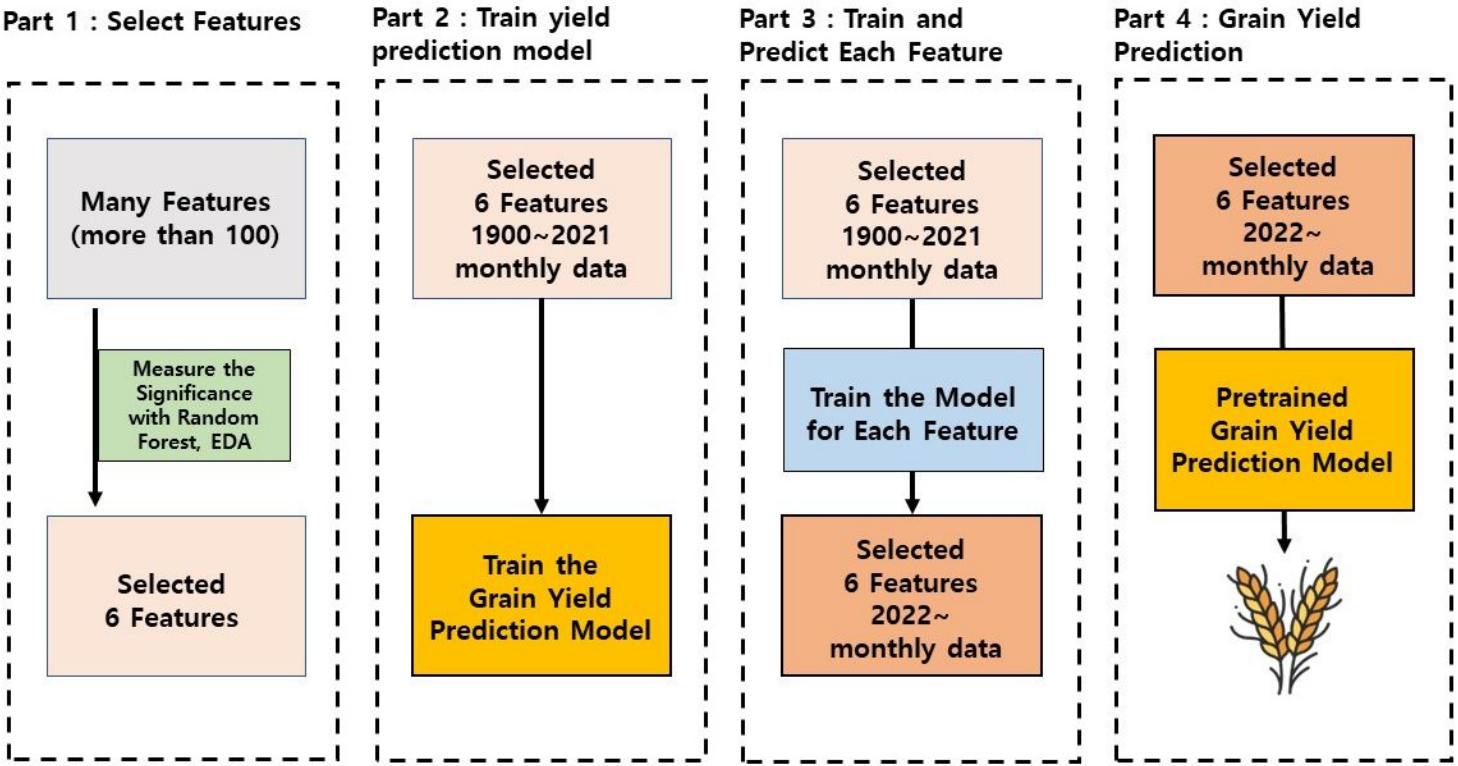
1. Grain production prediction model (Part 2)
2. 6 major features prediction data (Part 1, 3)

By combining those result,

we can predict future grain production!

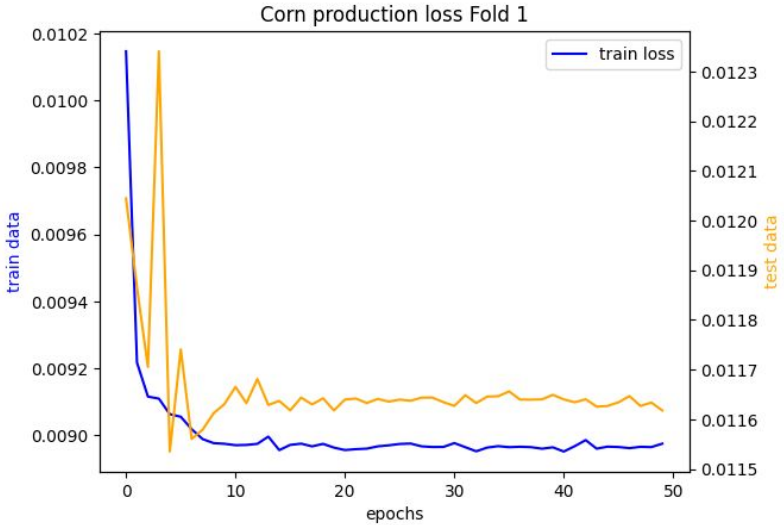


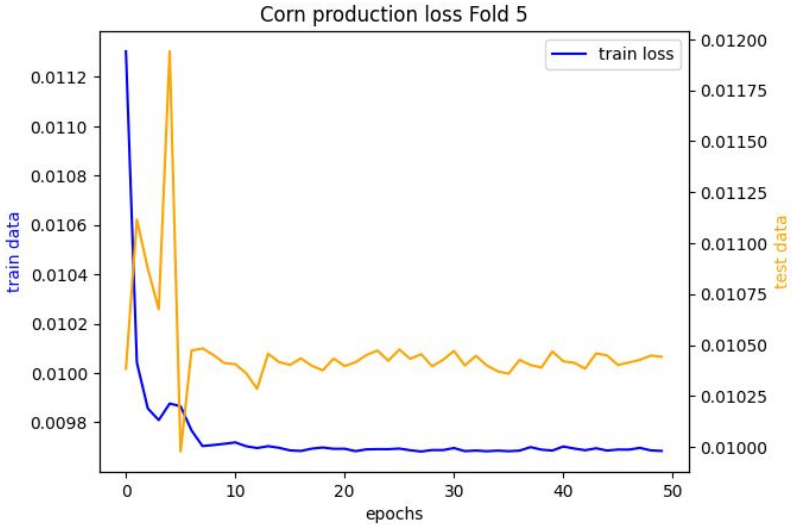
Model (Whole view)



Performance (K-Fold)

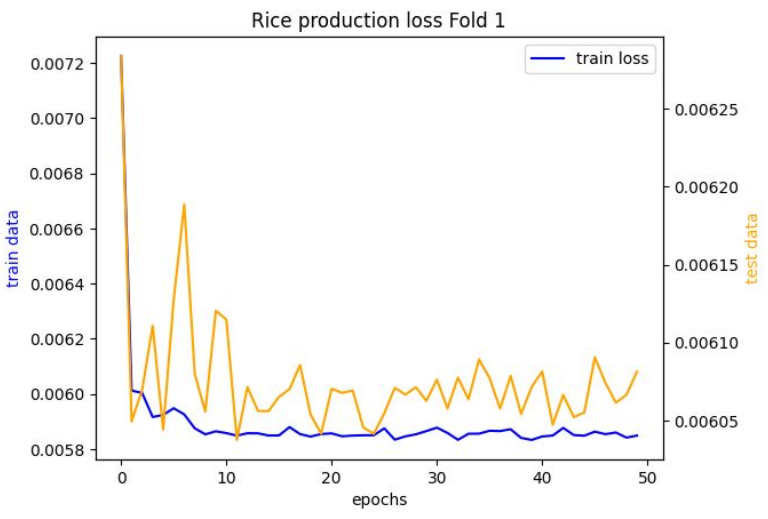
Corn: Test loss converged at 0.0105 ~ 0.0116

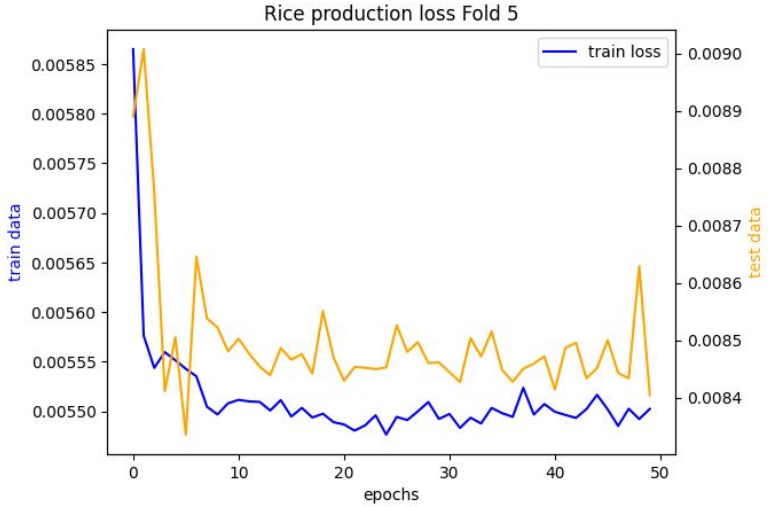




Performance (K-Fold)

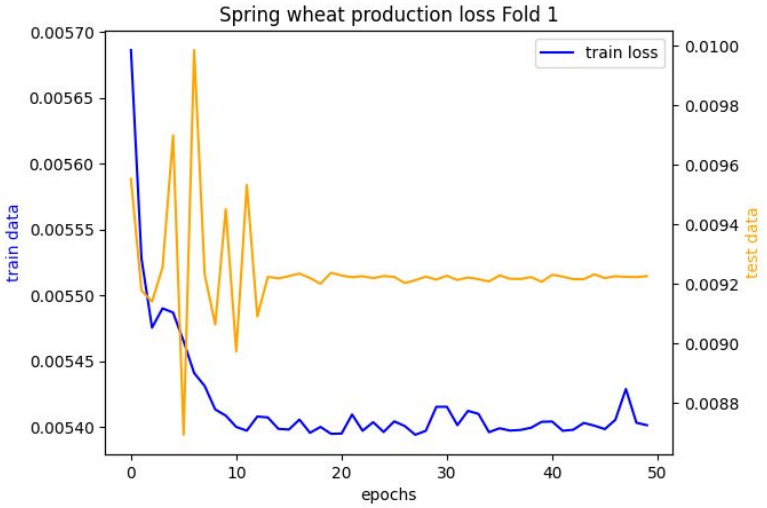
Rice: Test loss converged at 0.006 ~ 0.0084

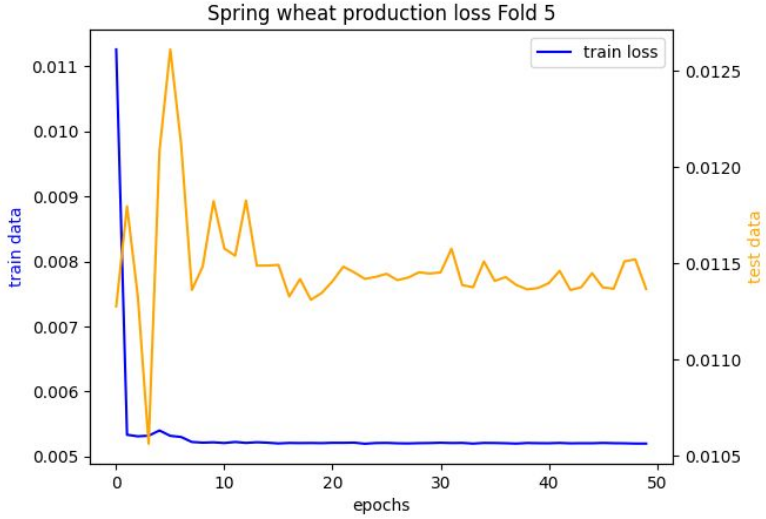




Performance (K-Fold)

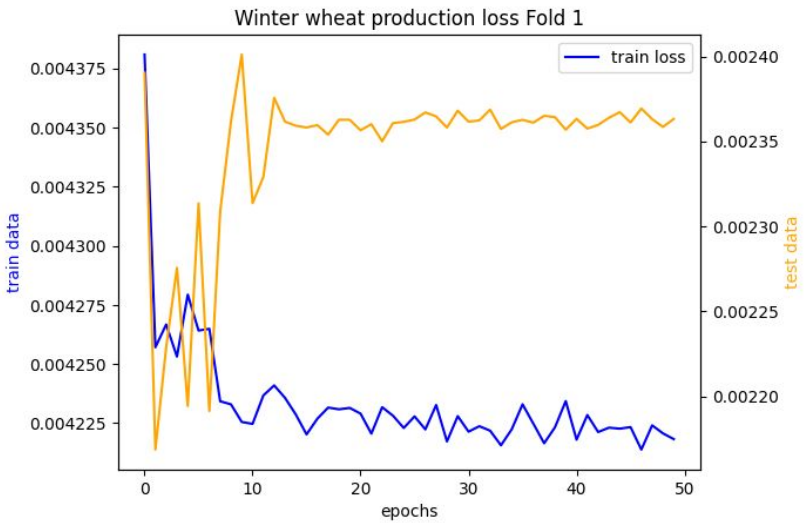
Spring wheat: Test loss converged at 0.0092 ~ 0.0115

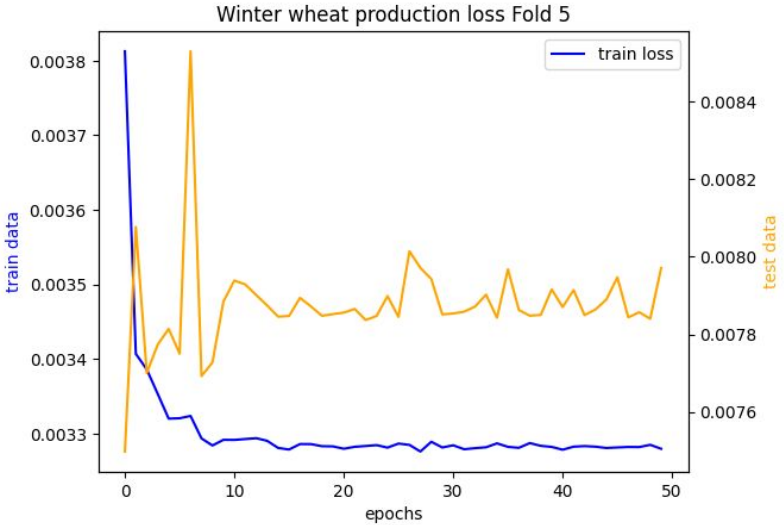




Performance (K-Fold)

Winter wheat: Test loss converged at 0.0035 ~ 0.0043





Performance (Prediction)

Winter wheat & corn production will be increased

Spring wheat & rice production will be decreased

Predictions	Spring Wheat	Winter Wheat	Rice	Corn
2022	2861981	1336457	1319915	25701286
2027	2739279 (-4.2%)	1357848 (+1.6%)	1272858 (-3.5%)	26001388 (+1.1%)
2031	2725797 (-4.7%)	1384889 (+3.6%)	1255904 (-4.8%)	26207812 (+1.9%)

Analysis

Features

- 6 major features - Population, CPI, food consumption, sunlight, co2 PPM, oil price
- Grain production highly correlated with economic & consumer features
    - CPI, population, food consumption, oil price
- Climate related features
    - sunlight, co2 PPM
    - why precipitation is not important?
        - modern cultivation has highly advanced irrigation system

Analysis

Prediction

- Each model’s test loss converged at 0.0035 ~ 0.001
- Corn and winter wheat production is expected to increase
- Rice and spring what production is expected to decrease

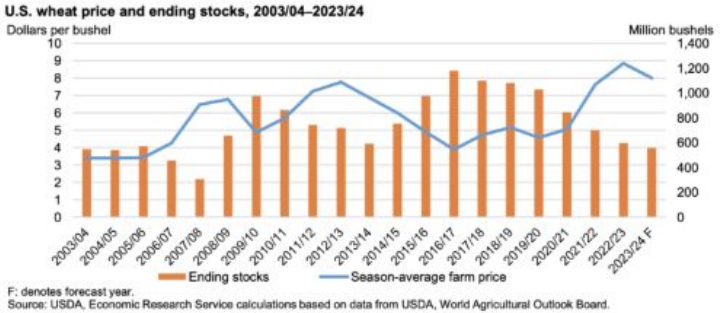
Analysis (Performance)

Source: USDA (U.S department of agriculture)

Feed Grain Supply Is Higher on Large Crops and Beginning Stocks

Expectations for larger crops places the forecast for 2023/24 U.S. feed grain production at 402 million metric tons, up from 358 million in 2022/23. Total supply of feed grains is projected at 443 million metric tons in 2023/24, which outpaces expected demand at 383 million metric tons. Forecast ending stocks for feed grains is projected at 59 million metric tons, up 21 million metric tons from the previous marketing year. Growth in the supply of feed grains is driven largely by corn. Corn supply in 2023/24 is forecast at 16,707 million bushels, up from 15,147 estimated in 2022/23. A larger corn crop and higher beginning stocks combine to bolster the corn supply.

Analysis (Performance)



Sowell, Andrew and Bryn Swearingen. Wheat Outlook: May 2023, WHS-23e, U.S. Department of Agriculture, Economic Research Service, May 16, 2023.

In related news, Bloomberg writer Michael Hirtzer reported yesterday that, "**It's so dry in Kansas that wheat plants are turning blue**.

In its annual Winter Wheat and Canola Seedings report, the USDA estimated the area planted to winter wheat for harvest this year at 36.95 million acres, up 11% from 2022 and the highest since 39.681 million acres in 2015. That figure was above the average of analysts' pre-report estimates, 34.485 million acres, and above the full range of guesses from 33.38 million to 36.2 million acres.

It seems our predictions are right at corn & wheat

Analysis (Performance)

1. Cannot explain the decreasing prediction of rice
    a. Maybe we can add more features related to rice production

2. Paper from NASA predict corn production will be decreased contrast to wheat prediction due to the climate change (affecting cultivation area)
    a. We didn't select the cultivation area in feature selection cause we think it's too obvious
    b. Maybe that makes the difference

Thank you.

Final presentation - Team 11
First Semester 2023, Artificial Intelligence Project (SWE3032-41)