**Team 5**

**Conceptual Proposal**

| Name | Contribution |
| --- | --- |
| Khoi | Proposal editing |
| Manuel | Powerpoint presentation |
| Liz Carillo | Github |
| Adam | Datasets and Gaussian regression modeling |

(Khoi’s part )

**Goal**:

**State the research question or hypothesis.**

Where is the best place to grow corn in the US in the next 10 years, using ML prediction on climate changes

**Explain the motivation behind this question, mentioning any relevant background information about why this topic or dataset is meaningful.**

Farming has always been the cornerstone of every society. In farming, weather rules the game, it dictates the yield rate of production. In the USA, corn is the most export and spearhead of agriculture. Using the tools that we have learnt in Machine Learning, we can predict the change of climate in the internal regions of the country where it best fit to grow corn, which can help to make a meaningful preparation for farmers and business in the future

(Adam’s part )

**2. Detailed Solution Plan**

As referenced in this report, https://www.nature.com/articles/s41612-020-00148-5, we can use a proven method machine learning model Gaussian Process Regression (<https://www.geeksforgeeks.org/gaussian-process-regression-gpr/>) to help us train and predict the ongoing climate changes.

**Import Libraries**

We need to import the libraries we’ll be using in the model creation.

* Pandas
* Numpy
* Matplotlib.pyplot - visualizations
* sklearn and its machine learning libraries

**Load dataset**

For the datasets, there is a free sourcing database at the national center for environmental information website-<https://www.ncei.noaa.gov/cdo-web/datasets> - where data sets in excess of gigabytes can be taken in csv formats over periods of time in regions, countries, states etc.

Once we have the dedicated region and dataset, we can download to a csv file and use pandas to read the data into a data frame using :

df = pd.read\_csv(“”).

**Exploratory Data Analysis**

With the datasets loaded we can then perform EDA, or exploratory data analysis, on the weather data. This includes checking and getting rid of missing numbers, checking basic visualizations with different plots, and getting the features and number of entries.

**Data Processing**

We can pad any missing values that we deem necessary and select to keep or get rid of features most needed in our goal for climate prediction.

After processing the data, we build the model with x and y variables (X\_train, X\_test, Y\_train, Y\_test) for testing and training in a .2(test) to .8(train) ratio.

**Define and Train the GPR Model**

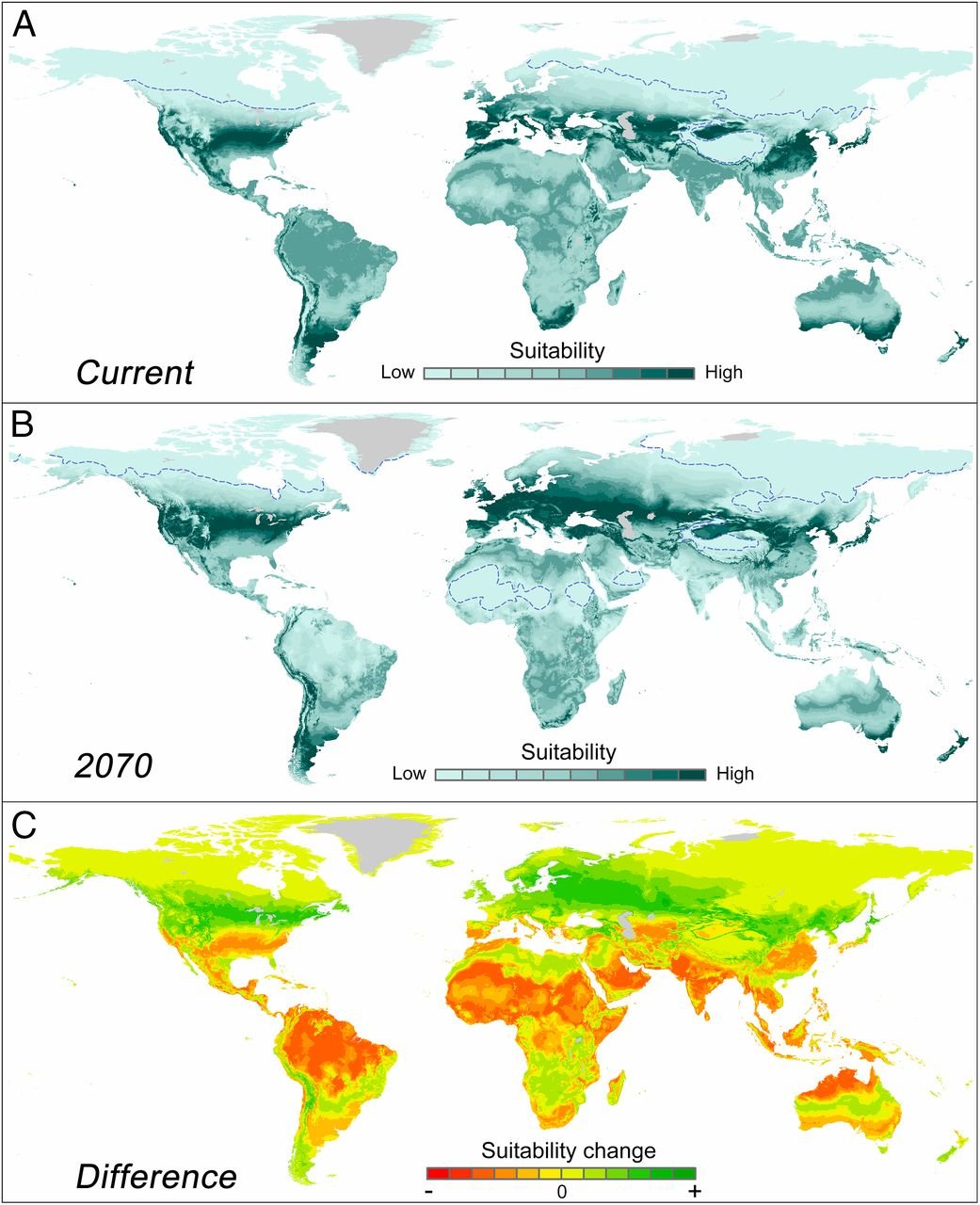
Using the variables with input data (X\_train, X\_test, Y\_train, Y\_test), we can fit those into the GPR (Gaussian Processing Regression) model to train and test.

**Make predictions and evaluate model**

With a complete trained and tested model on historical data of any region and timeline that deem noteworthy, we can now start predictions to figure out the trend of climate change. Charting visualizations of past data and comparing future predictions visualizations, we can get a clearer picture of what direction we are headed towards, whether it be generally colder, warmer, or stay sporadic.

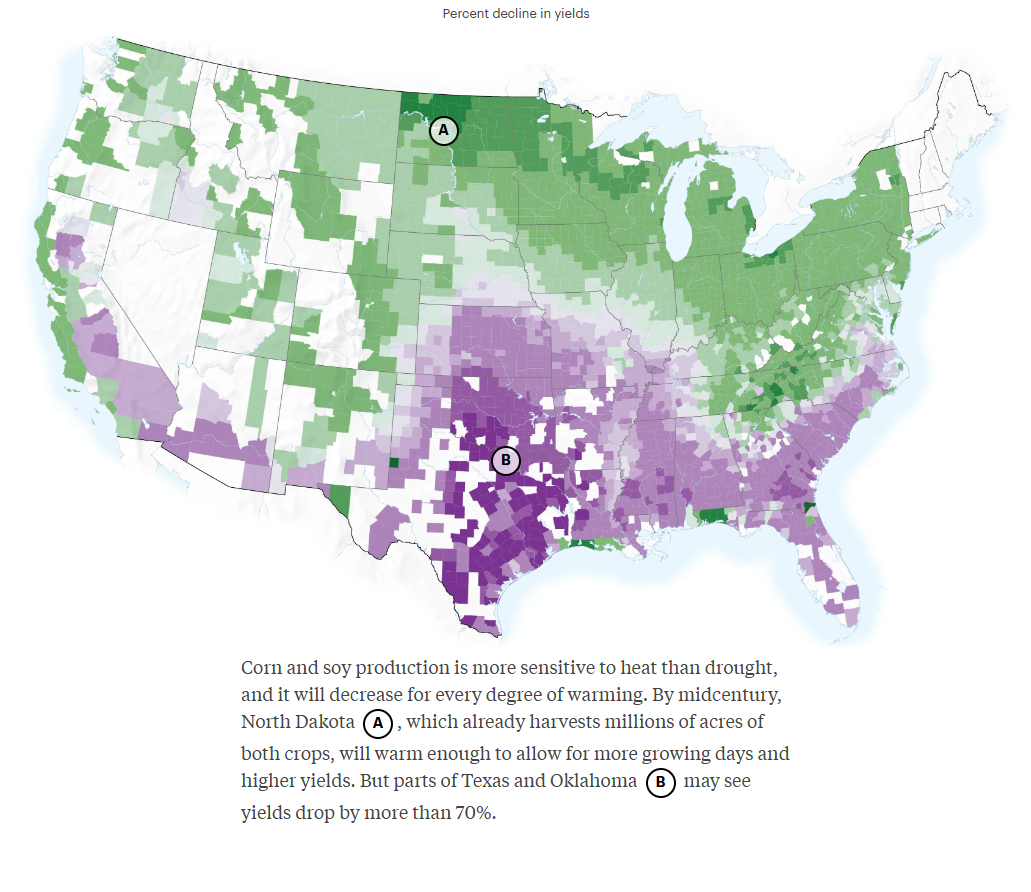
**3. Test Plan**

The model is intended to obtain accurate predictions about climate change based on the indicators in the dataset. For example, the template shared by Adam shows the results with the test model. Explain what was done in the test model and what results were obtained



*Source: https://www.pnas.org/*

As can be predicted using Gaussian Process Regression, the temperature suitable for growing crops, globally, will move north, by 2070. We are hoping to achieve a similar result as displayed by the heat map.



*Source:* [*https://projects.propublica.org/climate-migration/*](https://projects.propublica.org/climate-migration/)

**Conclusion**:

In short, prepare to move to the North!

It can be seen from climates change reports and predictions of global warming going to affect the yield crops across the entire USA. As temperature changes gradually over time in the south, so come the changes in human migration for better agriculture in the US.

By knowing this fact, it serves businesses and farmers tremendously to prepare for their future move, as farming has always and will always be the cornerstone of human civilization.