

# Drought Prediction Using Machine Learning

Devbrat Hariyani, Computer Science

Mentor: Dr. Ariane Middel

School of Computing and Augmented Intelligence



## Problem

Droughts devastate the:

1. Economy - cost the U.S. billions of dollars (energy-related)
2. Ecology – widespread degradation of ecosystems

## Research Goal

Develop a machine learning algorithm that:

- maintains a high accuracy in predictions
- is efficient and scalable
- is able to visualize data

The study uses Random Forests with data from the U.S. Drought Monitor.

## Why Random Forests?

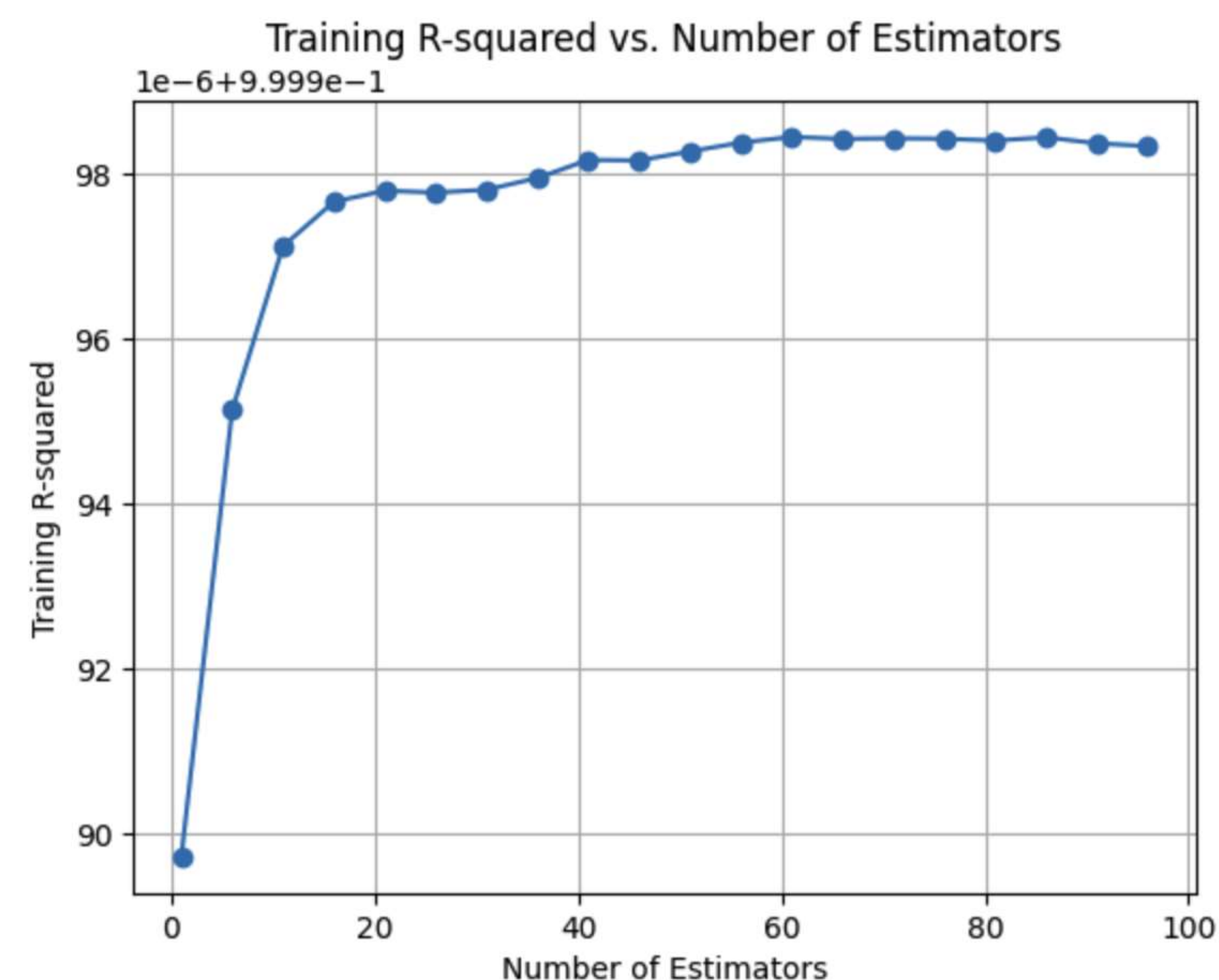
1. High Accuracy: runs efficiently on large datasets<sup>[1]</sup>
2. Estimates Missing Data: maintains accuracy when a large proportion of the data is missing<sup>[2]</sup>

## Dataset Used

- The U.S. Drought Monitor (USDM) maps drought areas using categories from D1 to D4 intensity.
- It includes abnormally dry (D0) regions and drought-free areas (None).

StateAbbreviation	None	D0	D1	D2	D3	D4	ValidStart	ValidEnd
AZ	8.17	91.83	55.78	32.15	6.09	0.00	2023-10-31	2023-11-06
AZ	8.11	91.89	55.82	32.15	6.09	0.00	2023-10-24	2023-10-30
AZ	8.11	91.89	54.04	30.36	6.09	0.00	2023-10-17	2023-10-23

## Results



Graph illustrates how the performance of the Random Forest model changes as the number of trees in the forest increases.

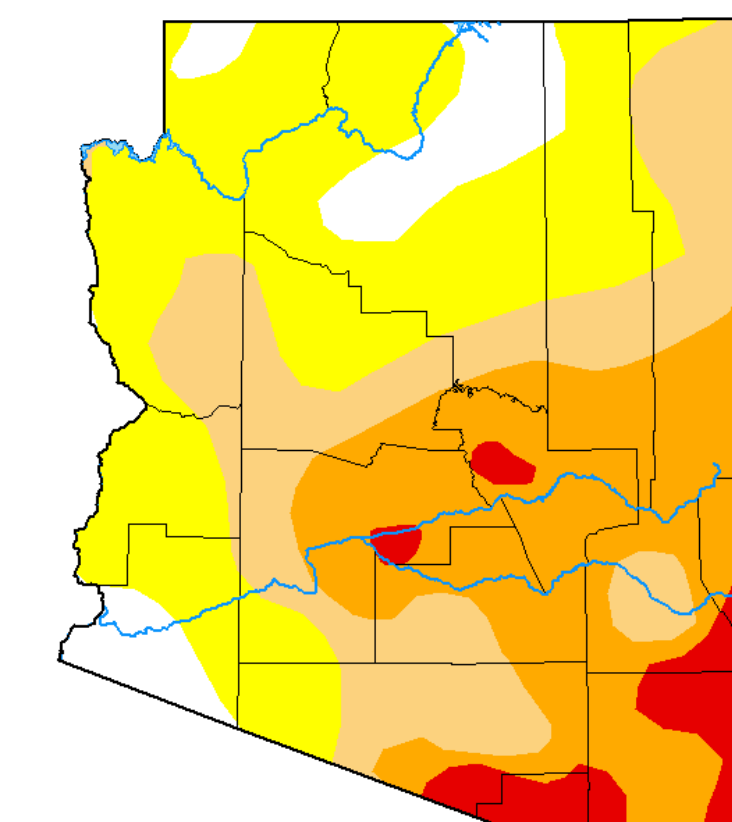
### References:

- [1] Park, H.; Kim, K.; Lee, D.k. Prediction of Severe Drought Area Based on Random Forest: Using Satellite Image and Topography Data. *Water* 2019, 11, 705. <https://doi.org/10.3390/w11040705>
- [2] Hobeichi, Sanaa, et al. "Toward a robust, impact-based, predictive drought metric." *Water Resources Research* 58.2 (2022): e2021WR031829.
- [3] Chao, Z., Pu, F., Yin, Y., Han, B., & Chen, X. (2018). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Journal of Sensors*, 2018, 1–9. 679 <https://doi.org/10.1155/2018/6184713>

## Libraries Used



## Future Work



- Correlate with spatial data (GIS) from the USDM for better predictions<sup>[3]</sup>

- Scale the data to the entirety of the United States
- Correlate with NDVI and other indices used in remote sensing<sup>[2]</sup>