**TEAM 6 Predicting Groundwater Levels in Spokane, Washington**

Marcos Ortiz, Meredith Sargent, Riti Bahl, Chelsea Gary, and Anireju Emmanuel Dudun

Github, Web App

**Objectives:**

- Create a machine learning model to forecast groundwater levels using information about surface water and weather.
- Create a web interface that allows stakeholders to explore model predictions to aid in planning and policy discussions.

**Data:**

- Groundwater level data, from selected wells located in Columbia Plateau basaltic-rock and basin-fill aquifers: one of the principal aquifers of the U.S.[1]
- Surface water data, from the Spokane River (along which our selected wells are located).[2]
- Precipitation data from Spokane, WA.[3]
- Weather data from Spokane County, WA.[4]

**Data processing:**

- Organized dataset from different sources into a format that allowed them to be used together.
- Engineered daily data from hourly weather and groundwater level data.
- Added a lagged precipitation feature upon comparing various delays to account for a delay between precipitation and impact on groundwater levels.
- Narrowed our feature list from 17 to 11 features that seemed likely to have the greatest predictive influence based on calculating the Pearson correlation coefficient.

**Models:**

- A baseline model to evaluate future model performance.
- Supervised learning models:
    - Ordinary Least Squares (OLS) linear regression
    - Convolutional Neural Network (CNN)
    - Recurrent Neural Network (RNN)
        - Implemented via a Long Short Term Memory model (LSTM)
- The neural networks were wrapped inside custom Scikit-Learn estimator/transformer classes so that hyperparameters could be tuned using a grid search.

**Web Interface:**

- Designed a web app using Streamlit.
- Allows stakeholders to access our model predictions and results interactively.
- Currently, our web app imports the pickled outputs from externally run python notebooks.

**Results:**

- Across each well, all models outperform the Baseline.
- For wells APK309 and AEK201, CNN outperforms the OLS linear regression.
- LSTM performs the best across all wells, with an RMSE between 0.34 - 2.3 ft.

**Implications:**

- This forecast can be used for planning and policy making related to water use policy.
- Can aid in detection of illegal overdrawing of critical groundwater resources.

---

[1] Source: Washington State Department of Ecology (WDOE)
[2] Source: United States Geological Survey (USGS)
[3] Source: National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI)
[4] Source: Openweather.com