**Week 3 Task Sheet: Advanced Modeling and Analysis**

**Objective:**  
By the end of this week, you should have moved beyond your initial feedforward network. You will explore models that better capture temporal or spatial patterns, interpret model predictions, and systematically evaluate your results with respect to the engineered climate and yield features.

1. **Data Preparation & Feature Engineering**

* Concatenate all available climate months for Alabama (Raven) and SC(Calvin).
* Engineer required features for each county and each month:
  + Daily/Monthly max, min, avg temp (convert K to °C), total precipitation, GDD, hot/cold days, drought/heatwaves.
  + Aggregate to one vector per county per month.
* For each county, create a sequence of these monthly vectors for 2022.

2. Merge with USDA Yield Data

* Join the engineered sequence (per county) to the county’s 2022 corn yield.

3. Format Data for Sequence Models

* X: 12-month (or growing season month for each crop) sequence per county, each with engineered features.
* y: Yield per county (target).

4. LSTM/GRU Modeling

* Split into train 8(0%)/test (20%) sets.
* Train LSTM and GRU models (with the same architecture).
* Report MAE, RMSE, and R² for both.

**2. Model Tuning and Validation**

* Perform hyperparameter tuning on your best-performing model. Tune learning rate, batch size, hidden layer size, and number of epochs using validation data.
* Use 5-fold cross-validation, if your data volume allows, to check model stability and to avoid overfitting—especially since you have limited counties and a single year.
* Track performance metrics (MAE, RMSE, R²) for each fold and average them.

**3. Error Analysis and Interpretation**

* For each county, compare actual vs. predicted yield (scatter plot or bar chart). Identify counties where predictions are consistently poor and list possible reasons (e.g., outlier climate, missing data, unmodeled soil effects).
* Use SHAP or feature importance methods (from tools like SHAP, permutation importance, or built-in PyTorch/TensorFlow functionality) to see which climate features are driving your model’s predictions.
* Perform an ablation study: drop one engineered feature at a time, retrain the model, and note the change in performance. Which features matter most?

**4. Incorporate Extreme Weather Events as Features**

* Specifically analyze whether years/counties with extreme hot days or drought periods (as calculated in Week 2) correspond to higher prediction errors.
* Add binary or count features for “extreme event occurred” and retrain the model to see if prediction accuracy improves.

**5. Document and Visualize Results**

* Plot cross-validation performance (bar plot of R²/MAE across folds).
* Visualize feature importance results.
* Create a summary table comparing the baseline model (Week 2) and advanced models (Week 3) on key metrics.

**Deliverables by End of Week 3:**

1. **Jupyter notebook(s)** showing code for advanced model(s), hyperparameter tuning, and feature importance analysis.
2. **Performance summary table** and at least three plots:
   * Actual vs. predicted yields by county
   * Cross-validation results
   * Feature importances/SHAP summary
3. **Short error analysis**: Identify counties/conditions where the model struggles and propose possible explanations.
4. **One-page reflection** on model improvements, what worked, what did not, and your next steps for Week 4 (e.g. ensemble models, or testing on other crops/months).