Predicting Water Quality Events in the Lynnhaven River Watershed

Faculty Member: Matthew Morena, Department of Mathematics
Community Partner: Lynnhaven River NOW, 3712 Holland Rd, Virginia Beach, VA 23452
Collaborating Research Partner: Cristin Pullman, Community Outreach Manager, LRNow

Project Summary

This project establishes a community partnership with Lynnhaven River NOW (LRNow), a local nonprofit dedicated to protecting and restoring the Lynnhaven River watershed in Virginia Beach. The Lynnhaven River is the principal waterway of Virginia beach and was once renowned for its thriving oyster populations. The river now suffers from decades of degradation caused by nutrient runoff, bacterial contamination, and urban development. LRNow regularly collects extensive water quality data, but needs additional support in order to detect meaningful trends and predict high-risk pollution events so that they can further optimize their outreach efforts. This project addresses these needs by using machine learning (ML) techniques to assess whether nutrient and bacteria levels correlate with seasonal cycles and rainfall patterns. A key hypothesis is whether runoff from lawn fertilizer is a primary contributor to the river's contamination. This research is explicitly use-inspired: it aims not only to generate new analytical insight, but also to produce practical tools and strategies LRNow can use to optimize their outreach, advocacy, and ecological restoration efforts.

Objectives and Methodologies

LRNow has provided ten years' worth of nutrient and bacteria data (2015–2024) from multiple sites within the Lynnhaven River watershed. After cleaning and processing the data, the project use a variety of ML techniques to identify meaningful patterns and develop predictive tools. The main goals are:

- Determine how contamination levels have changed over time and identify contamination hot spots.
- Determine how pollutant levels respond to rainfall and seasonal patterns using time-series models like ARIMA.
- Use multivariate regression to determine which environmental factors (e.g., rainfall, temperature, time of year) most strongly influence spikes in nutrient and bacteria levels.
- Build supervised ML models (e.g., Random Forest, CatBoost) to predict and classify high-risk contamination events.
- Use clustering methods (e.g., K-means, DBSCAN) to identify spatial locations with similar contamination patterns, helping LRNow to optimize their restoration effortsprioritize

With these objectives, LRNow will be able to predict when and where contamination is likely to occur, which improves their ability to design more effective restoration and outreach efforts.

Expected Outcomes and Broader Impacts

At the end of the project, LRNow will have a much clearer, data-driven understanding of when, where, and under what conditions contamination is most likely to occur in the Lynnhaven watershed. This includes statistical models to support ongoing monitoring, as well as predictive tools that can flag high-risk scenarios based on different environmental conditions. LRNow is especially interested in testing the hypothesis that lawn fertilizer runoff is a key driver of watershed contamination. If that turns out to be true, LRNow will an opportunity to enhance their outreach of responsible fertilizer use. This is something LRNow is already working on, but would now be able to back with local data.

The project will also develop practical, interactive maps and dashboards that LRNow can use to better communicate results to the community. More broadly, the project has potential to positively impact local public health by identifying and reducing waterborne illness risks. The project's insights could also help local urban planners balance sustainable economic development with environmental concerns.