# **Domain Application CA1 Notes**

**INTRODUCTION & GOALS**

We are going to try to determine if there is a way to accurately predict floods based on factors like time of year, weather, and location.

Dataset is based on historic station data across Bangladesh from 1948 to 2013.

Data found is obviously Bangladesh, but our hope is that if we can accurately predict data for Bangladesh that similar models could be potentially applied to other countries *e.g. what happened in Midleton last year.*

**ETHICS**

Inaccurate predictions can lead to ineffective or harmful interventions, eroding trust in analytics *i.e. the results could lead to poor decisions being made⇒reduced trust in predictions ⇒ safety concerns in our case.*

Predictive models might fail to account for the disproportionate impact of climate disasters on marginalised communities *i.e. poorer areas might flood more because of bad infrastructure etc.*

**BUSINESS VALUE**

Potential for predicting or at least determining high contributing factors to flood risk.

**INTRODUCTION**

Climate change is a topic that has been debated and discussed almost daily for decades, across media, throughout industry, and even in the day-to-day lives of the populus. While the majority of Irish people love to vent their frustrations about the terrible weather and how much we long for a bright and warm summer, the people of Ireland know first-hand how drastic the impact of climate change can be, especially with regard to the effect of climate change on natural disasters. Take for example the catastrophic damages done to localities in Cork in the latter months of 2023 caused by flooding following Storm Babet, ([RTE report](https://www.rte.ie/brainstorm/2024/0229/1435149-midleton-storm-babet-oct-2023-flooding-attribution-study-climate-change/) / [Midleton Study](https://www.worldweatherattribution.org/climate-change-made-the-extreme-rainfall-associated-with-flooding-in-midleton-ireland-more-likely-and-more-intense/#:~:text=On%20the%2017th%20and%2018th,months%20of%20above%20average%20rainfall.)) where months of heavy rainfall culminated in intense downpours covering the entire town in up to nearly six feet of water in some areas. While floods are not uncommon in the area, they are becoming more frequent and more intense.

Research from NASA ([NASA](https://www.earthobservatory.nasa.gov/images/148866/research-shows-more-people-living-in-floodplains)) shows that “the proportion of the world’s population exposed to floods grew by 20 to 24 percent” between 2000 and 2021. One of the regions noted to have a high rate of flood exposure was Bangladesh, where NASA claimed that roughly 27 million people were exposed as of July 2007. The population of Bangladesh during 2007 was roughly 144 million people meaning that 18.75 percent of the nation’s population were exposed to flooding during that period. An average of 21 percent of the country is flooded each year, largely in part due to heavy rainfall during monsoon season but mostly due to catchment areas beyond the country’s borders, where an estimated 80 percent of the floodwater flows from the river Ganges, the Brahmaputra-Jamuna, and the Meghna into Bangladesh. In 1998, more than 75 percent of the total area of the country was flooded through a combination of heavy rainfall in the country and in the surrounding catchment areas as well as an unfortunate synchronisation of peak flows coming from the major rivers. While Bangladesh is obviously a country of great concern with regard to the dangers of flooding, it is not the only country facing this problem. The problem is of course global, as shown by the events in the south of Ireland in 2023, and the findings of NASA’s research which are not even fully up to date as their research was released at the end of 2021.

Prior work has been undertaken in this field with the aim of predicting flood occurrences. The paper of Antwi-Agyakwa et al. [REF] discusses a range of flood prediction tools, how they have evolved as the field of research has expanded, and the shortcomings of many of the tools as well. In this paper they discuss the use of machine learning models such as tree-based models, Support Vector Machines, and Ensemble Prediction Systems. Motta et al. [REF] sought to deliver a spatial prediction tool for flood events in Lisbon to provide decision makers with a higher level of information to implement proactive measures and to allow emergency and recovery services to be optimally prepared. This paper discusses the use of methods like logistic regression, Support Vector Machines, Gaussian-Naïve Bayes, Random Forest, K-Nearest Neighbours, and Multi-Layer Perceptrons to classify whether a flood will take place. Both of these papers serve as a strong stepping stone for this project as they provide a great insight into the machine learning methods that have already been applied in this area of research.

**GOALS**

With the impending threat of greater and more dangerous flooding looming on the horizon, this paper aims to determine whether predictive analytics can be effectively applied to determine the likelihood of a flood occurring. To evaluate this, we will use a historic dataset consisting of weather station data across Bangladesh ranging from 1948 to 2013. Using this data, we hope to train a machine learning model that will provide robust and accurate predictions with relation to flooding occurring. By building this model we hope to demonstrate that flooding can be predicted and in doing so damages to infrastructure and threats to the population can be mitigated, and in some cases flooding may potentially be prevented completely. It is important to acknowledge that there is of course an array of human factors that can influence flooding such as poor infrastructure as well as some wholly unpredictable events that can contribute to or cause flooding, however we aim to deliver a baseline for prediction based on weather data.

**REFERENCES**

[1]K. T. Antwi-Agyakwa, M. K. Afenyo, and D. B. Angnuureng, “Know to Predict, Forecast to Warn: A Review of Flood Risk Prediction Tools,” Water, vol. 15, no. 3, p. 427, Jan. 2023, doi: <https://doi.org/10.3390/w15030427>.

‌ [1]M. Motta, M. de Castro Neto, and P. Sarmento, “A mixed approach for urban flood prediction using Machine Learning and GIS,” International Journal of Disaster Risk Reduction, vol. 56, p. 102154, Apr. 2021, doi: https://doi.org/10.1016/j.ijdrr.2021.102154.

‌

# **CA2 Notes**

k-NN model use for Bangladesh: <https://ieeexplore.ieee.org/abstract/document/9331199>

Flooding ML models built using SKLearn: <https://ieeexplore.ieee.org/abstract/document/9800023>

Case Study of Kebbi State Nigeria flooding (using SKLearn, NumPy, etc.): <https://ieeexplore.ieee.org/abstract/document/9800023>

KNN for rapid flood forecasting: <https://www.researchgate.net/publication/364250025_A_rapid_forecast_method_for_the_process_of_flash_flood_based_on_hydrodynamic_model_and_KNN_algorithm>