# Air Quality Prediction Using Machine Learning in Jordan

By:
**Maryam Imad Aljanabi**
**201725027**

Supervisor**:**
**Dr. Mohammad Shkoukani**

Co-Supervisor:

**Dr. Mohammad Hijjawi**

Master Thesis

Submitted in Partial Fulfillment of the Requirements for the Master Degree

in

Computer Science

Deanship of Research & Graduate Studies

Applied Science Private University

Amman – Jordan

January 2020

## ABSTRACT

Air pollution is one of the most vital hazards to human health nowadays, it is an invisible killer that takes a lot of human lives every year. The aim of this study is to

develop an air quality prediction system using machine learning techniques in Amman, Jordan and specifically in King Al-Hussein Public Parks area. Due to time importance when it comes to large datasets and computing resources, this thesis focused on reducing the number of features which led to a major improvement in time as well as the results. We compared the Artificial Neural Network, Support Vector Regression, Extreme Gradient Boosting, and Decision Tree Regression on three dataset combinations, then performed a comprehensive feature selection step to reduce the amount of features required for prediction. We found that the Artificial Neural Network was the best performing algorithm which was used to build the prediction system.

The developed prediction system predicted the numerical concentrations of $O_3$, $NO_2$, $SO_2$, and $PM_{10}$ then categorized these numbers to either normal or dangerous. The time was greatly reduced from before and after the feature selection step by about 91% for $O_3$, 80% for $NO_2$, 92% for $SO_2$, and 90% for $PM_{10}$. The final system scored high $R^2$ values and low errors for all the four pollutants indicating its good performance and its low error rates as compared to the previous literature