

**EMAIL OPEN AND CLICK RECOMMENDATION
ENGINE FOR PEARSON NOTIFICATION PLATFORM**

Tharindu Dulaj Koggalahewa

A dissertation submitted in partial fulfilment of the requirement for the
Degree of MSc Big Data Analytics

Department of School of Computing

Informatics Institute of Technology, Sri Lanka

in Collaboration with

Robert Gordon University, Aberdeen, Scotland

2023

Abstract

This research presents the development of an email open and click status prediction application for Pearson's Notification Platform (PNP) using machine learning algorithms, with a focus on the email subject line as a predictor. The objective of this study was to design and implement a model that can accurately predict the probability of an email being opened and a link in the email being clicked by the recipient. The methodology involved collecting a dataset of PNP emails and their corresponding open and click statuses, and then applying various feature engineering techniques to extract relevant features for the prediction model. Several machine learning algorithms, including Logistic Regression, Support Vector Machines, Decision Trees, Random Forest, and Multilayer Perceptron, were implemented, and compared for their prediction accuracy. The results showed that the Support Vector Machines algorithm achieved an accuracy of 61% for predicting open status for users of existing products, while the Random Forest algorithm achieved an accuracy of 61% for predicting open status for users of new products. For click status prediction, the Multilayer Perceptron algorithm achieved an accuracy of 84% for users of existing products, and the Logistic Regression algorithm achieved an accuracy of 84% for users of new products.

The developed system can be integrated with the existing Pearson Notification Platform, allowing product owners, business analysts, campaign runners, and the sales force of Pearson to access prediction data before sending emails. The contribution of this research lies in the development of a practical and effective email open and click status prediction application that can be used by Pearson.

Keywords: Email Open Status, Email Click Status, Machine Learning, Pearson Notification Platform (PNP)