Sequence Analysis of Search Terms

A Major Qualifying Project completed in partial fulfillment of the requirements for the

Degree in Bachelor of Science at Worcester Polytechnic Institute, Worcester, MA

In

Computer Science

By

Ying Lu   _____

Nicholas Paganetti   _____

Xiaoren Yang   _____

Submitted to: Professor Mohamed Eltabakh

December 14, 2016

_____

ABSTRACT

This project was done in correspondence with Demandware, a company interested in improving aspects of consumer experiences for several established client companies. The project is about analyzing consumer behavior on merchandise websites through a large data set provided by Demandware, and developing tools to produce meaningful results for use in several analytical purposes. More specifically, the data used was related to consumer searches on the merchandise sites that Demandware has collected over several years, and several other statistics about the user's experience on the site. This information was explored in three main ways: analyzing single search terms, sequences of search terms, and the connection between a single search term and product. The results are presented with both a Graphical User Interface (GUI) and Application Programming Interface(API) to be used for demonstration and development purposes.

AUTHORSHIP

This report was written by Ying Lu, Nicholas Paganetti, and Xiaoren Yang. All chapters were done in collaboration and are the responsibility of the group.

CONFIDENTIALITY

The rest of the report is omitted due to the confidentiality of its content. However, the complete report was reviewed by our advisor Professor Mohamed Eltabakh and is submitted and deposited in Demandware Inc.