SCIENTIFIC ADVISOR'S REFERENCE

Program:	Master in Business Analytics and Big Data
Student:	Gao Gan, Fu Songyuan, You Junwen
Title of thesis:	Development of a Website Classification Model for Quiet Media

Justification of the topic choice. Accuracy in defining the aim and objectives of the thesis. Justification of the topic choice; accuracy in defining the aim and tasks of the thesis; originality of the topic and the extent to which it was covered; alignment of the thesis' topic, aim and objectives.

The project for the thesis was initiated by the Quiet Media company, which is the developer for telecommunication operators. The project task for students was to provide technical improvement for the Quiet Media company in the process of providing services to telecommunication operators. Specifically, the goal of the master thesis is to build a high-performance text classification model for Quiet Media company to automatically detect sites with inappropriate content and to assess the performance of the model based on risk analysis from the business perspective view.

The developed model should help to improve the business process of advertising. It relies on machine learning and deep learning techniques to build a binary classification model.

Considering all the above, the master thesis topic seems to be justified. The title, research problem and research questions are clearly specified. The goal is feasible, and accurately expressed and the tasks are defined in a good form.

Structure and logic of the text flow. Logic of research; full scope of the thesis; alignment of thesis' structural parts, i.e. theoretical and empirical parts.

The thesis is organized logically. The paper consists of three chapters.

Structurally, the research is divided into four chapters. The first chapter provides an overview of the business-related background of the Quiet Media company, and the operation mechanism of the company is introduced. Research the goal, tasks, and objectives are defined as well. The second chapter provides a brief related literature review and analysis of the current state of the market and technologies for inappropriate site filtering. The third chapter is focused on the process of classification model development and evaluation. The flow of data gathering and preprocessing is also described in chapter three. The fourth chapter provides the analysis of the business performance of the model developed. The business recommendations for Quiet Media based on model evaluation and risk analysis from a business perspective view are also provided in chapter four.

The process of data analysis in this research is organized and employed according to the CRISP-DM methodology principles. Both theoretical and empirical parts of the thesis are consequently aligned in terms of their structure.

Quality of analytical approach and quality of offered solution to the research objectives. Adequacy of objectives coverage; ability to formulate and convey the research problem; ability to offer options for its solution; application of the latest trends in relevant research are for the set objectives.

Author demonstrates ability to formulate and convey the research problem and to offer options for its solution applying the latest trends in relevant research.

Quality of data gathering and description. Quality of selecting research tools and methods; data validity adequacy; adequacy of used data for chosen research tools and methods; completeness and relevance of the list of references.

The dataset was provided by Quiet Media but students should have collected the content of websites by themselves. The quality of selecting research tools and methods is good. The authors demonstrate the adequacy of used data for chosen research tools and methods. The data samples are described and analyzed. The list of references is relevant.

Scientific aspect of the thesis. Independent scientific thinking in solving the set problem/objectives; the extent to which the student contributed to selecting and justifying the research model (conceptual and/or quantitative), developing methodology/approach to set objectives.

The main scientific contribution of the thesis is the development of a machine learning solution for Quiet Media that can be used to detect websites with inappropriate content. The development methods were selected based on up-to-date relevant scientific papers as well as on study of experience with modern commercial solutions. Overall, the scientific contribution of the thesis is not outstanding but sufficient for the authors to be awarded the required degree.

Practical/applied nature of research. Extent to which the theoretical background is related to the international or Russian managerial practice; development of applied recommendations; justification and interpretation of the empirical/applied results.
The main practical contribution of the thesis is the development of the classification model that can be used for the detection of websites with inappropriate content by Quiet Media. For the developed website classification model, a financial model was built and a risk assessment was performed. The business recommendations for Quiet Media company regarding the developed website classification model deployment are provided.
Quality of thesis layout. Layout fulfils the requirements of the Regulations for master thesis preparation and defence, correct layout of tables, figures, references.
Layout fulfils the requirements of the Regulations for master thesis preparation and defence. All the tables and figures are properly edited and represented. There are some misspellings, typos, and formatting issues in the text that do not affect the quality of the paper as a whole.
Originality of the text. All sources of match identified by the Safe Assign system follow the allowed cases, the paper does not contain any elements of plagiarism.
All sources of match identified by the Safe Assign system follow the allowed cases, the paper does not contain any elements of plagiarism.

The Master thesis of Gao Gan, Fu Songyuan and You Junwen meets the requirements for master thesis of MiBA program thus the author of the thesis can be awarded the required degree.

Date: 08.06.2022

Scientific Advisor:

Olga Tushkanova
Candidate of Technical Sciences
Graduate School of Management,
St. Petersburg State University