**The Impact of Code Review Sentiment on Pull Request Success: A Comparative Analysis of Selected Pull Requests on GitHub**

**Author: Tayebe Kavousi Document: Thesis Outline**

[1. **Introduction**](#_Toc190951780)

[1.1. Background and Motivation](#_Toc190951781)

[1.2. Research Objectives](#_Toc190951782)

[1.3. Research Questions](#_Toc190951783)

[1.4. Scope and Limitations](#_Toc190951784)

[1.5. Research Contributions](#_Toc190951785)

[1.6. Thesis Structure](#_Toc190951786)

[**2.** **Related Work**](#_Toc190951787)

[2.1. Overview of Existing Studies](#_Toc190951788)

[2.1.1. Prior Research](#_Toc190951789)

[2.1.2. Research Gaps & Limitations](#_Toc190951790)

[2.2. Terminology and Definitions](#_Toc190951791)

[2.2.1. Sentiment analysis](#_Toc190951792)

* [Definition](#_Toc190951793)
* [Application](#_Toc190951794)
* [Challenges](#_Toc190951795)

[2.2.2. Deep Learning Models](#_Toc190951796)

* [Transformer Models](#_Toc190951797)
* [DistilBERT, DeBERTa, CodeBERT](#_Toc190951798)
* [Fine-tuning](#_Toc190951799)
* [Model Performance](#_Toc190951800)

[2.2.3. Code Review and PR Success](#_Toc190951801)

* [Definition and Importance](#_Toc190951802)
* [PR Success Metrics](#_Toc190951803)
* [Factors Affecting PR Acceptance](#_Toc190951804)

[2.2.4. Statistical Methods](#_Toc190951805)

* [Chi-Square Test](#_Toc190951806)
* [Cramér’s V](#_Toc190951807)

[**3.** **Methodology**](#_Toc190951809)

[3.1. Research Approach](#_Toc190951810)

[3.2. Model Selection](#_Toc190951811)

[3.2.1. Training Set](#_Toc190951812)

[3.2.2. Preprocessing data](#_Toc190951813)

[3.2.3. Fine-Tuning](#_Toc190951814)

[3.2.4. Model Evaluation](#_Toc190951815)

[3.3. Data Collection](#_Toc190951816)

[3.3.1. Sample Size Calculation](#_Toc190951817)

[3.3.2. Data Collection Criteria](#_Toc190951818)

[3.3.3. Data Cleaning](#_Toc190951819)

[3.3.4. Collected Information](#_Toc190951820)

[3.4. Sentiment classification](#_Toc190951821)

[3.4.1. Comments and Reviews Sentiment Classification](#_Toc190951822)

[3.4.2. PRs Sentiment Measurement](#_Toc190951823)

[3.4.3. PR Success Metrics](#_Toc190951824)

[3.5. Hypothesis Testing](#_Toc190951825)

[3.6. Validity and Reliability](#_Toc190951826)

[**4.** **Results and Findings**](#_Toc190951827)

[4.1. Labelled Data Preparation](#_Toc190951828)

[4.1.1. Dataset Balancing](#_Toc190951829)

[4.1.2. Data Cleaning and Preprocessing](#_Toc190951830)

[4.1.3. Data Splitting and Sampling Strategy](#_Toc190951831)

[4.2. Sentiment Analysis Model Development](#_Toc190951832)

[4.2.1. Model Selection](#_Toc190951833)

[4.2.2. Training and Fine-Tuning](#_Toc190951834)

[4.2.3. Baseline Comparison](#_Toc190951835)

[4.2.4. Performance Metrics and Error Analysis](#_Toc190951836)

[4.3. PR Data Collection](#_Toc190951837)

[4.3.1. Retrieving PRs, Comments, and Reviews](#_Toc190951838)

[4.3.4. Data Processing and Preparation](#_Toc190951841)

[4.4. Sentiment Analysis on PR Data](#_Toc190951842)

[4.4.1. Sentiment Classification of PR Comments and Reviews](#_Toc190951843)

[4.4.2. Aggregated PR Sentiment Scores](#_Toc190951844)

[4.5. Statistical Analysis and Hypothesis Testing](#_Toc190951846)

[4.5.1. PR Success Measurement](#_Toc190951847)

[4.5.2. Hypothesis Testing](#_Toc190951848)

[4.5.3. Findings for Research Question #1](#_Toc190951849)

[4.5.4. Findings for Research Question #2](#_Toc190951850)

[**5.** **Discussion**](#_Toc190951852)

[5.1. Interpretation of Results](#_Toc190951853)

[5.1.1. Model Performance Analysis](#_Toc190951854)

[5.1.2. Impact of Sentiment on PR Success](#_Toc190951855)

[5.3. Discussion of Limitations](#_Toc190951857)

[5.3.1. Limitations in Sentiment Classification Models](#_Toc190951858)

[5.3.2. Uncontrolled External Factors Affecting PR Success](#_Toc190951859)

[**6.** **Conclusion**](#_Toc190951861)

[6.1. Summary](#_Toc190951862)

[6.2. Recommendations for Future Research](#_Toc190951863)