وزارة التعليم العالي والبحث العلمي
جامعة الموصل
كلية علوم الحاسوب والرياضيات
قسم البرمجيات

تصنيف متطلبات البرمجيات باستخدام التعلم العميق

رسالة مقدمة
الى مجلس كلية علوم الحاسوب والرياضيات في جامعة الموصل
كجزء من متطلبات نيل شهادة ماجستير علوم في
البرمجيات

من قبل

عبدالرحمن مصطفى علي

بإشراف

أ.م.د ندى نعمت سليم

1444 هـ

2023 م

المستخلص

تعتبر هندسة المتطلبات اول واهم مرحلة في تطوير وبناء البرمجيات, اذ تهدف الى وصف وتوضيح متطلبات النظام, واحد المهام المهمة في هندسة البرمجيات هي تصنيف متطلبات البرمجيات كمتطلبات وظيفية ومتطلبات غير وظيفية والذي يساهم في تنظيم ودقة متطلبات البرمجيات واكمال النظام بشكل صحيح, تحدد المتطلبات الوظيفية مسؤولية ووظيفة النظام بينما تمثل المتطلبات غير الوظيفية عوامل جودة البرمجيات, والتمييز بين هذه الفئات يمثل احد التحديات التي ترام مهندسو البرمجيات, عادة ما يتم كتابة ووصف المتطلبات باستخدام اللغات الطبيعية وهذا قد يتسبب في ظهور العديد من العيوب والمشاكل نتيجة الغموض والتناقض في استخدام اوصاف مختلفة لنفس المتطلب وعدم وجود هيكلية منتظمة.

يهدف البحث الى بناء نظام مقترح لتصنيف متطلبات البرمجيات آليا باستخدام تقنيات وميزات التعلم العميق الى اصنافها الوظيفية والغير الوظيفية بفئاتها لمساعدة مهندسي البرمجيات في اختيار فئة او صنف المتطلب.

النظام المقترح يتكون من نموذجين, النموذج الاول هو نموذج التصنيف الثنائي والذي يتم من خلاله تصنيف المتطلبات الى المتطلبات الوظيفية والمتطلبات غير الوظيفية, اما النموذج الثاني فهو نموذج التصنيف المتعدد يتم خلاله تصنيف متطلبات البرمجيات الى المتطلبات الوظيفية وعشر فئات من المتطلبات غير الوظيفية .

تم بناء النموذجين المقترحين باستخدام تقنيات وميزات الشبكات العصبية التلافيفية(CNN) Convolutional Neural Networks وتمثيل وتضمين النصوص باستخدام نموذج word2vec بعد معالجتها باستخدام تقنيات معالجة اللغات الطبيعية, واستخلاص الصفات والميزات المهمة.

تم تطبيق وتنفيذ نماذج النظام المقترح على مجموعة البيانات المستحدثة في البحث والناتجة من دمج مجموعة البيانات (PROMISE) ومجموعة البيانات (nfr), والتي تمثل متطلبات البرمجيات المكتوبة باستخدام اللغات الطبيعية والممثلة بصيغة نصية, حيث تم الحصول على (1271) متطلبا مقسما الى (11) فئة.

تم تقسيم البيانات الى (90%) للتدريب و(10%) للاختبار والتأكد من صحة النموذج, حققت دقة التدريب لنموذج التصنيف الثنائي(99.44%) ودقة نموذج التصنيف المتعدد (96.25%), اما دقة الاختبار فبلغت في نموذج التصنيف الثنائي (98%) ونموذج التصنيف المتعدد(97.62%).

Ministry of Higher Education and
Scientific Research
University of Mosul
College of Computer Science and
Mathematics
Department of Software

Software Requirements Classification using Deep Learning

**A Thesis Submitted to the Council of the College of
Computer Science and Mathematics
University of Mosul
as a Partial Fulfillment of Requirements
for the Degree of Master of Science
in Software**

Abdulrhman Mustafa Ali

Supervised by

Ass .Prof .Dr. Nada N. Saleem

2023 A.D. **1444 A.H.**

Abstract

Functional requirements define the responsibility and function of the system, while non-functional requirements represent software quality factors, distinguishing between these classes is one of the challenges facing software engineers. Requirements are usually written and described using natural languages, and this may cause many defects and problems to appear as a result of ambiguity and contradiction in the use of different descriptions of the same requirement and the lack of a regular structure.

Intelligent technologies, especially machine learning and deep learning, have been used in scientific and engineering fields, and this has led to increased accuracy and reduced time, effort, and cost in all operations and various fields of application.

The research aims to build a proposed system automatically classify software requirements using deep learning techniques and features into functional and non-functional Classes in order to help software engineers choose a Class or Classes of requirement.

The proposed system consists of two models, the binary classification model, through which requirements are classified into functional requirements and non-functional requirements. And the Multiple Classification Model, through which software requirements are classified into functional requirements and ten categories of non-functional requirements.

The two proposed models were built using the techniques and features of Convolutional Neural Networks (CNN) and text representation and embedding using the (word2vec) model after processing it using natural language processing techniques, and extracting important characteristics and properties.

The proposed system models were applied and implemented on the dataset developed in the research resulting from the merger of the PROMISE dataset and the nfr dataset, which represents the requirements for programs written using natural languages and represented in text. appearance. Where (1271) requirements were obtained, divided into (11) chapters.

The data was divided into (90%) for training and (10%) for testing and ensuring the validity of the model. The training accuracy of the binary classification model was achieved (99.44%) and the accuracy of the

multiple classification model was (96.25%). As for the accuracy of the test, it reached (98%) in the binary classification model and (97.62%) in the multiple classification model.

The results obtained showed the accuracy and completion of the proposed models using CNN deep learning architecture with the word2vec method, as they achieved the required results and were efficient and able to successfully classify script requirements with higher accuracy than previous research.