

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ
ФАКУЛЬТЕТ КОМПЮТЕРНО-ІНФОРМАЦІЙНИХ СИСТЕМ
І ПРОГРАМНОЇ ІНЖЕНЕРІЇ
КАФЕДРА ПРОГРАМНОЇ ІНЖЕНЕРІЇ

БАРАНСЬКИЙ МАКСИМ ОЛЕКСАНДРОВИЧ

УДК 004

**РОЗРОБКА СИСТЕМИ МАШИННОГО ПЕРЕКЛАДУ НА ОСНОВІ
НЕЙРОМЕРЕЖЕВИХ ТЕХНОГОЛІЙ З ВИКОРИСТАННЯМ ВЕКТОРА
МЕТРИК ЯКОСТІ**

121 «Інженерія програмного забезпечення»

Автореферат

дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль
2019

Роботу виконано на кафедрі програмної інженерії Тернопільського національного технічного університету імені Івана Пулюя Міністерства освіти і науки України

Керівник роботи: Доктор технічних наук, професор
Пастух Олег Анатолійович,
Тернопільський національний технічний університет
імені Івана Пулюя,

Рецензент: кандидат технічних наук, доцент
Дмитроца Леся Павлівна,
Тернопільський національний технічний університет
імені Івана Пулюя,

Захист відбудеться __ _____ 2019 р. о ___ годині на засіданні екзаменаційної комісії №1 у Тернопільському національному технічному університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. Руська, 56, навчальний корпус №1, ауд.

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми роботи. Розробка системи машинного перекладу на основі нейромережових технологій з використанням вектора метрик якості на сьогоднішній день досить актуальна, оскільки вона дозволяє вияснити яка з архітектур є кращою для машинного перекладу та наскільки є точний переклад.

Мета роботи: Розробка двох систем машинного перекладу на основі різних нейромережових технологій з використанням вектора метрик якості та зробити порівняння цих архітектур.

Об'єкт, методи та джерела дослідження. Нейромережові архітектури рекурентна нейронна мережа (RNN) та архітектура «Довга короткочасна пам'ять» (LSTM).

Наукова новизна отриманих результатів: розроблено системи машинного перекладу на основі нейромережових технологій з використанням вектора метрик якості, метрики якої не використовуються у онлайн перекладачів.

Практичне значення отриманих результатів.

Отримано нові програмні системи машинного перекладу з використанням двох різних нейромережових технологій, та з обчисленням вектора метрики якості для визначення точності перекладу, здійсненими цими системами. Ці системи є кросплатформеними, так як використовувалася скриптова мова програмування.

Апробація. Окремі результати роботи доповідались на VII Науково-технічній конференції «Інформаційні моделі, системи та технологій». Тернопіль, ТНТУ, 11-12 грудня 2019 року.

Структура роботи. Робота складається з розрахунково-пояснювальної записки та графічної частини. Розрахунково-пояснювальна записка складається з вступу, 7 частин, висновків, переліку посилань та додатків. Обсяг роботи: розрахунково-пояснювальна записка – _____. формату А4.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі проведено аналіз проблематики дослідження функцій зі змінним періодом рядами Фур'є та створення відповідних інформаційних технологій їх аналізу.

В розділі «Розробка програмної системи» було розроблено, протестовано та верифіковано інформаційну систему машинного перекладу на основі нейромережових технологій з використанням вектора метрик якості.

В розділі «Спеціальна частині» було розглянуто саме поняття нейромережових технологій та дві технологій, які були використанні у рамках цієї магістерської, а також про метрики якості машинного перекладу та метрику METEOR, яка також була використана в даній роботі.

В розділі «Обґрунтування економічної ефективності» розглянуто питання розрахунку економічної ефективності і терміну окупності капітальних вкладень.

В розділі «Охорона праці та безпека в надзвичайних ситуаціях» проаналізовано поняття про виробничий травматизм та професійні захворювання.

Було досліджено профілактику травматизму та захворювань. Також було проаналізовано питання безпеки в надзвичайних ситуаціях, таких, як фактори ризику і можливі порушення здоров'я користувачів комп'ютерної мережі та планування заходів цивільного захисту на об'єкті у випадку надзвичайних ситуацій захисту.

У загальних висновках щодо дипломної роботи було розглянуто що таке машинний переклад, які технології машинного перекладу використовуються, зокрема і нейромереві. Було розроблено дві системи машинного перекладу з різними архітектурами та за допомогою вектора метрик якості було визначено, що машинний переклад з використанням LSTM дає кращі результати.

В додатках до пояснювальної записки знаходиться короткий лістинг програми.

ВИСНОВКИ

В дипломній роботі магістра було розроблено інформаційну систему дослідження електрокардіограм на основі функцій зі змінним періодом. Було розглянуто проблему дослідження таких функцій, .

У даній інформаційній системі були розроблені шаблони сценаріїв. За допомогою цих сценаріїв користувач має можливість імпортувати структуровані дані записів електрокардіограм, обрати необхідний діапазон значень для дослідження, після чого отримати результати як в текстовому так і в графічному вигляді.

В ході розробки дипломної роботи магістра були отримані результати, які показують, що при правильному виборі та впровадженні інформаційної системи, можна отримати суттєві конкурентні переваги і можливості для зростання, за допомогою пришвидшення дослідження.

Розроблена інформаційна система містить в собі всі необхідні функціональні можливості для імпорту, обробки та відображення результатів роботи. Також інформаційна система володіє необхідним рівнем захисту від несанкціонованих модифікацій вихідного коду. Таким чином, розроблена інформаційна система доводить свою життєздатність, ефективність і необхідність її впровадження.

СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

1. Баранський М.О. Розробка системи машинного перекладу на основі нейромеревих технологій з використанням вектора метрик якості / Баранський М. О. // Збірник тез доповідей VII Науково-технічної конференції «Інформаційні моделі, системи та технологій», 11-12 грудня 2019 року. — Т. : ТНТУ, 2019.— С. 48. — (Інформаційні моделі, системи та технологій).

2. Палій В.М. Комп'ютерна програма «Аналіз RR-інтервалів та оцінка варіабельності серцевого ритму» / Палій В.М. // Свідectво про реєстрацію

авторського права на твір, 7 червня 2018 року. — №79650 — Міністерство економічного розвитку і торгівлі України.

АНОТАЦІЯ

У дипломній роботі розроблено інформаційної системи дослідження електрокардіограм на основі функцій зі змінним періодом.

Проаналізовано літературні джерела, визначено поняття функцій зі змінним періодом а також розглянуто приклади таких функцій у повсякденному житті.

Метою дипломної роботи є розробка інформаційної системи дослідження електрокардіограм на основі функцій зі змінним періодом.

Дипломна робота присвячена розробці інформаційної системи дослідження електрокардіограм на основі функцій зі змінним періодом. Розроблено власний програмний продукт для імпорту, аналізу та відображення результатів опрацювання даних електрокардіограм.

Практичне значення одержаних результатів. Отримано нову інформаційну систему яка дозволяє швидко та зручно імпортувати, вибрати дані для аналізу та отримати результат згідно наперед визначеним алгоритмам опрацювання вхідних даних, яка представляє собою програмний продукт для операційних систем сімейства Windows

Ключові слова: АНАЛІЗ, ДОСЛІДЖЕННЯ, ДАНІ, СЕРЕДОВИЩЕ, ІНФОРМАЦІЙНА СИСТЕМА, ЕЛЕКТРОКАРДІОГРАМА, RR-ІНТЕРВАЛ, CSV.

ANNOTATION

In the thesis the information system of the research of electrocardiograms on the basis of functions with a variable period is developed.

The author analyzes literary sources, defines the concept of functions with a variable period, and also examines examples of such functions in everyday life.

The purpose of the thesis is to develop an information system for the study of electrocardiograms based on functions with a variable period.

The thesis is devoted to the development of an information system for the study of electrocardiograms based on functions with a variable period. The software for importing, analyzing and displaying the results of processing of electrocardiographic data is developed.

In the fourth section, the main principles of information systems development were considered, and the native analias of programming languages and developmental sinks and also the service for infographic were carried out. Visual Studio 2015 has been featured in the development industry since it is fast and easy to use and a free license is also provided for students.

The fifth section deals with the question of calculating the economic efficiency and the payback period of capital investments.

In the sixth section, the methods of environmental statistics, electricity extraction from the combustion of mineral fuel and environmental pollution in this case, as well as ways to reduce it. Internet resources were analyzed for the availability of additional information on the solution of the tasks.

The seventh section analyzes the concept of occupational injuries and occupational diseases. The prevention of injuries and diseases was investigated. Security issues were also analyzed in emergencies, such as risk factors and possible violations of the health of users of the computer network and planning of civil protection measures at the facility in the event of emergency situations of protection.

The practical value of the results. A new information system has been obtained that allows you to quickly and conveniently import, select data for analysis, and obtain a result according to predetermined input processing algorithms, which is a software product for Windows family operating systems.

Keywords: ANALYSIS, RESEARCH, DATA, ENVIRONMENT, INFORMATION SYSTEM, ELECTROCARDIOGRAM, RR-INTERVAL, CSV.