

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

КОСТАНОВИЧ РОМАН ВІКТОРОВИЧ

УДК 004.89

**СЕРВІС ДЛЯ РОЗПІЗНАВАННЯ КАПЧ ЗА ДОПОМОГОЮ НЕЙРОННИХ
МЕРЕЖ**

8.0501302 «Інженерія програмного забезпечення»

Автореферат

дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль 2017

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

Керівник роботи: доцент кафедри програмної інженерії
Михалик Дмитро Михайлович,
Тернопільський національний технічний університет
імені Івана Пулюя,

Рецензент: доктор технічних наук, професор
Лупенко Сергій Анатолійович,
Тернопільський національний технічний університет
імені Івана Пулюя,

Захист відбудеться 23 лютого 2017 р. о 9.⁰⁰ годині на засіданні екзаменаційної комісії №33 у Тернопільському національному технічному університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. вулиця Руська, 56, навчальний корпус №1, ауд. 101

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми роботи. Сервіс для розпізнавання капч за допомогою нейронних мереж допомагає тестувати системи захисту сайтів та додатків від автоматичних комп'ютерних програм, які імітують роботу реального користувача, дозволяє тренувати нейронні мережі для оптичного розпізнавання символів. Даний сервіс також є площадкою для купівлі і продажу послуги розпізнавання капч. Реалізовано всі необхідні функції щодо ефективності та надійності роботи. Програмне забезпечення дозволить користувачам пришвидшити та спростити процес розпізнавання тексту з зображень.

Мета роботи: Створити сервіс для розпізнавання капч з відкритим кодом. Зменшити і, в окремих випадках, відмовитися від ручної праці людини в процесі розпізнавання тексту з зображень шляхом використання нейронних мереж. Створити систему, яка дозволить оцінити систему захисту сайту чи додатку від автоматизованих програм, що імітують роботу реальних користувачів. Програмне забезпечення дозволить швидко і бездоганно виконання функцій пов'язаних з розпізнаванням тексту з зображень.

Об'єкт, методи та джерела дослідження. Основними об'єктами дослідження є процеси розпізнавання капч за допомогою спеціального сервісу і за допомогою нейронних мереж. Методом дослідження є математичне моделювання, та метод аналізу і синтезу для розбивання всього процесу на частини. Засобом для цього виступає комп'ютерна техніка і статистичні дані.

Наукова новизна отриманих результатів:

- виконано аналіз існуючих продуктів, їхні недоліки та переваги;
- проаналізовано, наскільки ефективно капча захищає сайт від автоматизованих програм;
- виявлено негативний вплив капч на користувачів сайту;
- проаналізовано переваги і недоліки розпізнавання капч людьми та програмами;
- наведено приклади простих і складних капч для систем розпізнавання;
- досліджено і проаналізовано аналогічні методи захисту сайту від автоматизованих програм;
- проаналізовано, як нейронні мережі допомагають в завданнях оптичного розпізнавання символів;
- розроблено сервіс для розпізнавання капч;
- розроблено модуль для розпізнавання деяких окремих видів капч.

Практичне значення отриманих результатів.

Розроблене рішення дозволить втілити переваги попередніх рішень та вирішити їхні недоліки. Програмний продукт являтиме собою загальне рішення для розпізнавання капч.

Апробація. Окремі результати роботи доповідались на Міжнародній науково-практичній конференції «Сучасні наукові інновації», Київ, Міжнародний центр наукових досліджень, 15-16 лютого 2017 р.

Структура роботи. Робота складається з розрахунково-пояснювальної записки та графічної частини. Розрахунково-пояснювальна записка складається з вступу, 5-ти частин, висновків, переліку використаних посилань та додатків. Обсяг роботи: розрахунково-пояснювальна записка – 97 аркушів формату А4, 2 додатки, графічна частина – 16 слайдів графічної частини.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У вступі описано основні терміни роботи і призначення сервісу для розпізнавання капч.

У розділі аналіз предметної області та використані технології дипломної роботи було описано предметну область, проблеми, з якими стикаються користувачі. Розглянуто основних користувачів системи, їхні ролі. Проаналізовано вибір баз даних, мови програмування та фреймворк, розглянуто архітектуру системи, описано основні використані технології. Розглянуто аналогічні готові продукти, виявлено їхні переваги та недоліки. Описано і проілюстровано код основних частин системи.

В розділі «Обґрунтування економічної ефективності» розглянуто два підходи до розробки даного продукту, а саме об'єктно-орієнтований і процедурний. Проаналізовано усі витрати і прибутки, потрібні складові для роботи, супровід і модернізацію програмного продукту та уточнено аналіз економічної ефективності вкладених інвестицій, обчислено різницю між двома підходами щодо витрат на виробництво і модернізацію.

В частині «Охорона праці та безпека в надзвичайних ситуаціях» розглянути норми охорони праці користувачів ПК, описано основні фактори які впливають на функціональний стан користувача ПК, вимоги ергономіки до організації робочого місця оператора ПК.

В розділі «Екологія» дано характеристику вимог до приміщень для експлуатації моніторів і ПЕОМ, проаналізовано шляхи дотримання цих вимог.

У загальних висновках щодо дипломної роботи описано використані в роботі програмні рішення, які забезпечують виконання поставленого завдання, зазначено оригінальні програмні рішення, прийняті автором в процесі роботи.

В додатках до пояснювальної записки додано технічне завдання і диск з програмою.

ВИСНОВКИ

Описано предметну область, та проблеми з якими стикаються її учасники. Розглянуто основних користувачів системи, їх ролі. Проаналізовано вибір бази даних, мови програмування та фреймворк, розглянуто архітектуру системи, описано основні технології, які використовувалися, а також застосований підхід до розробки бази даних.

Розроблено модель даних яка повністю задовольняє меті, розробки, описано основні клас, які відповідають предметній області. Наведені основні аргументи вибору бази даних, і інших ключових етапів розробки які суттєво впливають на процес розробки.

Описано основні принципи роботи класів в межах фреймворку Vert.x, зроблено обґрунтування вибору технологій для створення API. Створене API повністю задовольняє вибраній темі роботи і виконує поставлені функції розробки.

Виконано обґрунтування проведених досліджень та розраховано основні економічні показники, які виявилися більш економічно вигідними для об'єктно-орієнтованого підходу, ніж для процедурного, що є очікуваним.

Враховані вимоги до охорони праці для користувачів ПК, для забезпечення комфортних умов праці. Враховано фактори що впливають на оператора ПК при роботі. Описано ергономіка організації робочого місця оператора ПК, описано основні методи для розміщення елементів на столі.

Охарактеризовано вимоги до приміщень для експлуатації моніторів і ПЕОМ та проаналізовано шляхи дотримання цих вимог.

СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

1. Костанович Р. В. Сервіс для розпізнавання капч за допомогою нейронних мереж // Тези доповіді на Міжнародній науково-практичній конференції «Сучасні наукові інновації». – Київ, МЦНД, 2017.

АНОТАЦІЯ

Дипломна робота на тему «Сервіс для розпізнавання капч за допомогою нейронних мереж» Костановича Романа Вікторовича. — Тернопільський національний технічний університет імені Івана Пулюя, Факультет комп'ютерно-інформаційних систем і програмної інженерії, Кафедра програмної інженерії, група СПм-61 // Тернопіль, 2017.

С. – 97, рис. – 6, табл. – 4, слайд. – 16, додат. – 2, бібліогр. – 20.

Методом дипломної роботи є дослідження та розробка технології для тестування систем захисту сайтів та додатків від автоматичних комп'ютерних програм, які імітують роботу реального користувача, і тренування нейронних мереж для оптичного розпізнавання символів.

Методи розробки базуються на технологіях Java, Vert.x, Cassandra, Redis, Python, OpenCV та FANN.

Використана методика проектування інформаційних систем з модульною структурою та застосуванням сучасних технологій веб-програмування. Здійснений програмний продукт реалізовує базові функції обробки інформації для вибраної сфери.

Ключові слова: БАЗИ ДАНИХ, СЕРВЕР, CAPTCHA, OCR, API, WEBSOCKET, HIGH PERFORMANCE, НЕЙРОННІ МЕРЕЖІ, JAVA, VERT.X, CASSANDRA, REDIS, PYTHON, OPENCV, FANN.

ABSTRACT

Diploma work on theme «Service for captchas recognition by using neural networks» by student Kostanovych Roman Viktorovych. — Ternopil Ivan Pul'uj National Technical University, Faculty of Computer Information Systems and Software Engineering, Software engineering department, group SPm-61 // Ternopil, 2017.

Pages. – 97, pictures. – 6, tables. – 4, slides – 16, add. – 2, bibl. ref. – 20.

The method of the thesis is to research and develop technology for the testing of site and application defense systems from automated computer programs that simulate real user work and training neural networks for optical character recognition.

The methods of development based on Java, Vert.x, Cassandra, Redis, Python, OpenCV and FANN.

Used methods of designing information systems with a modular structure and modern technologies of web programming. Result product implements basic functions of processing information for the chosen area.

Keywords: DATABASES, SERVER, CAPTCHA, OCR, API, WEBSOCKET, HIGH PERFORMANCE, NEURAL NETWORKS, JAVA, VERT.X, CASSANDRA, REDIS, PYTHON, OPENCV, FANN.