Dang Nguyen Quang Huy

Data Scientist

PROFILE

- Male
- 0395049120
- hdang1696@gmail.com
- https://github.com/ZeusCoderBE
- https://www.linkedin.com/in/huyhocdata/
- Linh Trung ward,Thu Duc City ,Ho Chi Minh City

SKILLS

MAIN SKILL

Programming Languages : Python, Java, C#, OOP

EDA and Visualization: Python (numpy ,pandas, seaborn ,matplotlib),Data Preprocessing ,Power BI,Excel

Machine Learning/Deep Learning: Python (scikit-learn ,tensorflow ,PyTorch) ,Regression,Clustering, Classification,NLP (Sentiment Analysis ,Text Generation, Machine Translation),Kaggle,Google Colab

Collection: BeautifulSoup, SSIS

Data Modelling:SQL Server, Relational Database (SQL), SSAS

Mathematics: Probability and Statistics, Linear Algebra

Big Data : Apache Hadoop, Apache Hive

English: TOEIC 605 (8/2023)

Version Control:Git

OTHER SKILL

Streamlit, Flask framework, NET Framework,AWS Cloud,Ubuntu Desktop/ Server

STRENGTH

Hardworking,Teamwork, Planning,Creative

OBJECTIVE

After graduation, I aim to pursue a career as an AI Engineer or Data Scientist Specialist. My primary interest lies in processing diverse data types, with a particular passion for Natural Language Processing (NLP) and Generative AI. I'm excited about the endless possibilities in these fields and look forward to creating valuable products that address real-world challenges.

EDUCATION

Major: Data Engineering **2021 - 2025**

School: Ho Chi Minh City University of Technology and Education

GPA : 8.59

Certificate - Coursera: Machine Learning Specialization **(3/2024)**

Certificate - English: TOEIC 605 **(8/2023)**

PROJEACT

I.Project Name: Applying Artificial Neural Networks to Build Text Generation Models as Part of the Generative AI Problem : (3/2024 - 4/2024)

Source: *https://github.com/ZeusCoderBE/Next_word_predicting*

Team : 1 (Individual Project)

Description: *Text generation is a challenge in generative AI.*

- I developed a large language-based model for automatically generating text.I collected land law data from ***https://thuvienphapluat.vn/*** using BeautifulSoup and preprocessed the data using NLP techniques. This includes extracting sentences, determining the meaning of related phrases, building a collocation dictionary, and generating input sequences using the N-GRAM method.
- I implemented Deep Learning architecture with embedding layers and SimpleRNN, leveraging TensorFlow and Keras libraries for model development and evaluation. Performance evaluation includes data such as accuracy, precision, acquisition, and F1 score. I deployed the model to the website using flask framework.

II.Project Name: Applying artificial neural networks to build models to analyze customer emotions based on comments and evaluation serves for determination business-related trends: (3/2024 - 5/2024)

Source: *https://github.com/ZeusCoderBE/NLP-clustering-word--Vietnamese-Sentiment-Analysis*

Team : 1 (Individual Project)

Description: *Sentiment analysis acombined with text generation:*

- I developed a project to analyze customer sentiment from comments using KNN,CNN and LSTM models. I used BeautifulSoup to collect comments and reviews from ***https://www.thegioididong.com/***. I preprocessed the data using NLP techniques, including removing punctuation, stop words, and symbols, combining meaningful Vietnamese words, reformatting text, encoding words, and creating a corpus. For the CNN model, I used word2vec (skip-gram) for word embedding to capture the relationship between words and their context. In the LSTM model, I used an embedding layer to represent word relationships and visualize the embedded words as vectors in space.
- In addition, I rebuilt the text generation model to automatically create comments for users, integrating it with the emotion recognition model to analyze emotions when sending comments. Models were evaluated using accuracy, precision, recall, and F1 score. I have implemented a sentiment analysis model on a web application using Flask. This helps mobile world stores improve business processes and product quality, attract more customers, and increase revenue.

III.Project Name: Building The Recommender System through content filtering and collaborative filtering: (1/2024 - 3/2024)

Source: *https://github.com/ZeusCoderBE/Recommender-System*

Team : 1 (Individual Project)

Description: *I implemented two recommendation algorithms sush as Content Filtering and Collaborative Filtering.*

1. Content Filtering:

- I created a vector representation for each movie using TF- IDF (item profiles).
- I trained a ridge regression model for each user to learn the weights (user profiles).
- I used item profiles and user profiles to predict and recommend movie ratings

2. Collaborative Filtering:

- I utilized two approaches: item-item and user-user.
- I calculated cosine similarity between items or users.
- I implemented a KNN model by selecting K similar users/items to predict rating scores

AWARDS

OUTSTANDING STUDENT SCHOLARSHIP IN THE FIELD OF STUDY **2023 - 2024**

School: HCMUTE

That school year, I had the highest score in my major