Di WU

Machine learning engineer/ Data scientist

✉ di.wu@student.uclouvain.be

📍 Verte Voie 46/305, Louvain la Neuve

📅 Date of birth 12/03/1993

🔗 www.linkedin.com/in/di-wu-171804151/

🏳 Chinese

📇 Europe B driving license

📞 +32498713345

LANGUAGES

Chinese

Native

English

C1

French

A2

COMPUTER SKILLS

Python

3 years experience

SQL

3 years experience

Tensorflow

2 years experience

Javascript

2 years experience

R

2 years experience

Docker

1 year experience

EDUCATION

Master in data science engineering

Université catholique de Louvain - Louvain la Neuve, Belgium
- From September 2018 to August 2021

Main courses:
Advanced Algorithms for Optimization, Algorithms in Data Science, Artificial Intelligence, Automatic Control Theory, Bayesian Statistics, Data Mining, Databases, Information visualization, Machine Learning, Mining Patterns in Data

WORK EXPERIENCE

Data science intern

Deloitte Belgium - Brussels - From February 2021 to June 2021

Achievements:
● Extract text from images and pdf documents using OCR services
● Extract table from client's invoices using Azure's Form Recognizer and AWS's Textract, SageMaker
● ETL for results with text and bounding boxes and store on MongoDB
● Build and train the active learning models using GPT3 and BERT as well as rest API encapsulation

Technologies: Python, Tensorflow, Azure, AWS, Active learning, NLP, BERT, MongoDB, huggingface

NLP intern

AG Insurance - Brussels - From October 2020 to January 2021

Achievements:
● Develop chatbot service for Health insurance of My Global Benefits insurance service on Chatlayer and Azure
● Develop chatbot service for Employee booking system on Chatlayer and Azure
● ETL dialogue dataset in three main service languages
● Compare the performance of two platforms with K fold cross validation and confusion matrix

Technologies: Python, Chatlayer, Azure, NLP, Hypothesis testing, K fold cross validation

CERTEFICATES

Spark and Python for Big Data with PySpark -- Feb/2021
Udemy certificate for using Spark with Python, including Spark Streaming, Machine Learning, Spark 2.0 DataFrames

Deep Learning Specialization -- Sep/2020
Coursera Deep Learning Specialization including 5 courses about Convolutional Neural Networks, Recurrent Neural Networks, LSTMs, Transformers and so on.

The first prize for national robotics competition -- Dec/2014
Won first prize in the Challenge of Robot for the Elderly in The China RobotCup