

Python Programming - UCI - 100% Online Courses. Learn the Most In-Dei



Helena H. Lee · 3rd in

Oracle

NLP Data Scientist at Oracle (Contract)

San Francisco Bay Area · 500+ connections · Contact info

About

- Authorized to work for all employers in the United States
- Languages: Python (Most of the time; 4+ years' experience), SQL, R, MATLAB, Spark
- Tools and Technologies: Machine Learning (Scikit-learn), Natural Language Processing (Gensim, S

Featured

BEHT, thus reduced the model complexety by 2/x and improved the training and interence speed

Analytics Research Centre - Research Engineer (Python, PyTont, TensorFlow, ElasticSearch)

2018-2020

in the Joint presearch centre of Company Mellog Life Joint Page 10 (Python, PyTon)

2018-2020

- *Ci st be joint research centre of Carnegie Mellon University and Singapore Management University, en Researched and implemented data science and statistical approaches to solve AI-related problems, predictive modeling, natural language processing, and natural language generation (NLC).
 *Proficient in data exploration data representation, model development, and model development.
- Proficient in data exploration, data preparation, model development, and model deployment.
 As the first author, published 1 research paper at ACM DPH'19 and 1 demo paper at ACM WWW'20
- Fine-tuned generative pre-trained language model GPT-2 on 1 million cooking recipe documents
 Deployed an interactive website which allows conditional recipe generation in real-time at RecipeGPT.org
- receiving 1000+ visits across 18+ countries and 2 paper citations in 6 months after release

 Proposed and implemented the world's first model to identify recipes which are unhealthful for diabetics.

Compared different document classification models (NBBoW, Word2Vec, RNN) and achieved an F1 of 0.85

BrainCo - Algorithm Team Intern (Python, Scikit-learn)

BrainCo is a brain-machine interface start-up founded by Harvard scientists, located in Somerville, MA

Designed the proof-of-concept experiments to correlate people's attention level with brain wave signals

Acer is one of the top personal computer vendors in the world, headquartered in Taiwan.

Developed a computer vision model [MobileNet. a light-weighted convolutional neural network] using Tensorflow to assess the quality of diabetic retinopathy images with an accuracy of 39%

Helena H Lee Resume [lastest].pdf

Activity

518 followers

Interested

Helena H. commented

Experience



Data Scientist

Oracle · Contract
May 2020 – Present · 6 mos

San Francisco Bay Area

Oracle's marketing data science team integrates digital marketing solutions that help B2B a B2C marketers

- Contract role focusing on analyzing marketing and sales data, including but not limited to: writing SQL queries, building data engineering pipelines, improving data quality, and training machine learning models
- Trained multi-label classification models on documents in PyTorch to identify the interests potential customers. The training involves over 6 million textual paragraphs and more than classes.
- Compared the pros and cons of different deep neural networks ranging from RNN, LSTM,
 Transformer, and BERT, thus reduced the model complexity by 27x and improved the trainir and inference speed

see less

monster code

Data Scientist

Monster Code Corporation · Contract
May 2020 – Oct 2020 · 6 mos

San Francisco Bay Area



Research Engineer

Living Analytics Research Centre (LARC) \cdot Full-time 2018 – Mar 2020 \cdot 2 yrs

LARC is the joint research centre of Carnegie Mellon University and Singapore Managemen University.

- Researched and implemented data science and statistical approaches to solve Al-related problems, including regression, classification, natural language processing, and recommensystems.
- Proficient in data exploration, data preparation, model development, and model deployme
- As the first author, published 1 research paper at ACM DPH'19 and 1 demo paper at ACM WWW'20.
- Fine-tuned generative pre-trained language model GPT-2 on 1 million cooking recipe documents
- · Deployed an interactive website which allows conditional recipe generation in real-time at

RecipeGPT.org, receiving 1000+ visits across 18+ countries and 2 paper citations in 6 mont after release

- Proposed and implemented the world's first model to identify recipes which are unhealthf for diabetics
- Compared different document classification models (NBBoW, Word2Vec, RNN) and achievan F1 of 0.85

see less





Algorithm Team Intern

BrainCo · Internship

2018 · less than a year Greater Boston Area

BrainCo is a brain-machine interface start-up founded by Harvard scientists, located in Somerville, MA

- Designed the proof-of-concept experiments to correlate people's attention level with brai wave signals
- Recruited 20+ volunteers across teams within the company to participate in electrophysic (EEG) experiments

see less



Data Scientist on Kaggle

 $\text{Kaggle} \cdot \text{Freelance}$

2017 – 2018 · 1 yr

- Attended the Kaggle contest: Corporación Favorita Grocery Sales Forecasting
- Conducted feature engineering and modelling on time series sales data (Time, Store, Item Quantity)
- Final world ranking: 58/1675 (Top 4%); Won the Silver medal



Research & Development Intern

Acer · Internship

2017 · less than a year

Acer is one of the top personal computer vendors in the world, headquartered in Taiwan.

 Developed a computer vision model (MobileNet, a light-weighted convolutional neural network) using Tensorflow to assess the quality of diabetic retinopathy images with an accuracy of 93%

Show fewer experiences ^

Education



National Chiao Tung University

Master of Science (M.Sc.), Bioinformatics and Systems Biology 2016 – 2018



National Chiao Tung University

Bachelor of Science (B.Sc.), Biological Science and Technology 2012 – 2016



Zhejiang University

Exchange student, Biology, General 2014 – 2014

Licenses & certifications



AWS Fundamentals: Going Cloud-Native

Coursera
Issued Apr 2020 · No Expiration Date
Credential ID VZ9LTTSHCA4B

See credential

Skills & Endorsements

Python · 6



Endorsed by Albert Hsu, who is highly skilled at ...



Passed: LinkedIn Assessments

Natural Language Processing (NLP)

TensorFlow

Industry Knowledge

Data Mining · 3 Machine Learning · 2 Research **Statistics** Data Science · 1 Deep Learning · 1 **Artificial Intelligence (AI) Algorithms Data Visualization Computer Science** Data Analysis · 2 **Data Modeling Data Structures Data Collection Statistical Data Analysis Big Data** Extract, Transform, Load (ETL) Tools & Technologies $\mathbf{R} \cdot 4$ Matlab · 5 **Amazon Web Services (AWS)** SQL

Git **Oracle Database**

PyTorch Keras

Other Skills ②

Apache Spark

Show less ~

Accomplishments



Publications

RecipeGPT: Generative Pre-training Based Cooking Recipe Generation and Evaluation System.

- · Characterizing and Predicting Repeat Food Consumption Behavior for Just-in-Time Intervent
- Estimating Glycemic Impact of Cooking Recipes via Online Crowdsourcing and Machine Learn
- Age-related differences in frontoparietal activity underlying creativity for convergent and dive thinking.

Interests



Sheng Fu in

Founder and General Partner — Purple ... 3,049,196 followers



台美人求學求職合作社

767 members



Artificial Intelligence, Deep Learn...

358,516 members



Silicon Valley Startup, Mento

1,579 members



Chinese Institute of Engineers / U...

119 members



James Manyika 🛅

Senior Partner at McKinsey & Cor 27,460 followers

See all

Linked in

About

Accessibility

Community Guidelines

Careers

Privacy & Terms ✓

Ad Choices

Sales Solutions

Mobile

Talent Solutions

Marketing Solutions

Advertising

Small Business



Questions?

Visit our Help Ce



Manage your a

Go to your Settin

Safety Center

LinkedIn Corporation © 2020



Messaging