Quoc Pham Ba Cuong

Data Scientist
Ho Chi Minh City University of Science
(+84) 378 050 314
pbcquoc@gmail.com
github.com/pbcquoc

OVERVIEW

I'm a Data Scientist, Deep Learning Engineer, Big Data Engineer.

WORK EXPERIENCE

Lecture 2018 - Now

Lecture at VietAI, Non-profit Organization

- I teach Deep Learning for Vietnamese student. My topic is about CNN, LSTM, Attention Mechanism for Machine Translation.

Data Scientist 2017 - Now

Data Scientist at Zalo Group, VNG Corporation

- I build a large-scale machine learning system that store, process and detect relationships of over 80 million users. The system stores historical data about users location (gps, wi-fi) and their personal information. I extract edge attribute of the super graph (16 billion edges) about their distance in several days, and build a machine learning model using these features to classify their relationship. My model has 86% precision and 0.9 AUC.
- Our user is annoyed by irrelevant notification of their friend activity. I suggest and build a smart notification system that leverage the relationship and use PageRank to rank tie strength. Our system improves user engagement by 10%.
- I build a dashboard for our machine learning system that support AB testing and monitor various measurement. The system helps us reduce experiment time.

Data Scientist 2016 - 2017

Lead Recommender System at PayTV, FPT Telecom

- I build a movie recommender system to serve 600k users using PredictioIO, MLLib Spark, Scala. In detail, I implement Alternative Least Square for Collaborative Filtering approach and Single Value Decomposition for Content-Base approach, after that i combine these models using a linear model. My system improves performance to 500% request-per-second and Click-Through-Rate increase 40%.
- I proposal and build a music recommender system using Generative Adversarial Networks which is a novel deep learning model. The input to system is rely on only raw audio so songs can be suggested without any meta-data. The input is converted to MFCC features and pass to my model to learn representation, then i use cosine similarity for ranking songs. My system result is impressive than other based on meta-data.
- My company business grow, we need a system that must to be scalable, flexible, easy to integrate new model, support AB testing which is an well-known business metrics, and lastly, realtime adapt to user when they watch new movies or musics. I proposal an architecture which combine Spofity's recommender system and Netflix's recommender system, and lead 4 teammates to build the system.
- I use user-item logs of our log system to determine linear trendline and suggest to user.

- I make alert system for our system to track anomaly and email to related person in time.

SIDE PROJECT

Chat Bot

Research and build chat bot using seq2seq which is LSTM encoder that encode sequence to fixed vector and try to decode into answer. I crawl tweet and comments as question and answer to train my model.

Image Synthesis

Research and build DCGAN model which is a generative model then use it to learn image representation that can use to image search.

Realtime Deep Learning Model

Design architecture and build deep learning model (for localization, detection, classification) that can run on small cheap device that have less than 512MByte memory and slow processor.

EDUCATION

Bachelor of Science - Honor Program 2012-2016

Ho Chi Minh City University of Science

"Generate Image Caption using Convolution Neural Netword and Long-Short Term Memory"

Thesis advisor: MSc Tran Trung Kien

We implement a novel model to generate image description. Our approach is association of two most successful deep learning models in Computer Vision and Natural Language Modeling that known as Convolution Network and Long Short-Term Memory. Our result can help blind man to see the world.

ACADEMIC ACTIVITY

Deep Learning 2015-Now

Ho Chi Minh City University of Science

MSc Tran Trung Kien

Our group have learnt more about Supervised Learning based on Bengio's book. It includes mathematics optimization, understanding deep learning, having some experiments to find out deep learning model behavior.

ACHIEVEMENT

- 2016 The Best Thesis, Ho Chi Minh City University of Science
- 2017 8th/220, Message Prediction Competition
- 2017 19th/418, CIKM AnalytiCup 2017: Lazada Product Title Quality Challenge

- 2017 Top 1 Zalo Hackathon, We build chatbot for selling product
- 2018 Top 4 Entropy, Data Analytics Competition
- 2018 Top 2 Zalo AI Challenge, The First AI Competition in Viet Nam
- 2018 Top 1 Cinnamon Hackathon, Vietnamese Optical Character Recognition

SKILL

Base Knowledge	Deep Learning, Machine Learning Computer Vision, Natural Language Processing Data Analysic, Statistics, Computer Graphics
Programming	Python, Cuda C/C++, MATLAB, Lua, Java, HTML/CSS, C/C++, C#
Libraries	Tensorflow, Theano, Torch7, Scikit-Learn, Pandas
Languages	English

ONLINE COURSE

- 2014 Learning From Data, Taught by Caltech Prof. Yaser Abu-Mostafa
- 2016 Convolutional Neural Networks for Visual Recognition, Taught by Stanford Ph.D Andrej Karpathy
- 2016 Deep Learning for Natural Language Processing, Taught by Stanford Ph.D Richard Socher
- 2017 Introduction to NLP, Taught by Stanford Prof. Dan Jurafsky & Chris Manning
- 2017 Introduction to Probability - The Science of Uncertainty, Taught by Prof. John Tsitsiklis