

# Curriculum Vitae

April 2021

## Vu Anh

Data - AI Department  
VNDIRECT  
[anhv.ict91@gmail.com](mailto:anhv.ict91@gmail.com)

<https://www.vndirect.com.vn/>

1 Nguyen Thuong Hien, Hai Ba Trung  
Ha Noi, Viet Nam

## ❖ Education

### B.A., Software Engineer

2009-2014

Good Certificate

Hanoi University of Science and Technology, Viet Nam

## ❖ Experience

### Natural Language Processing Researcher

2016-present

2016-present: **underthesea**

underthesea

Open source Vietnamese NLP toolkit, 800+ github star

2020: Fourth Place in the shared task of Dependency Parsing VLSP'20

Archive 78.58 *UAS* and 70.04 *LAS*

2018: Second Place in the shared task of Sentiment Analysis VLSP'18

Archive 64.2  $F_1$  in restaurant domain, 58.7  $F_1$  in hotel domain

2016: Second Place in the shared task of Named Entity Recognition VLSP'16

Archive 85.6  $F_1$ , SOTA in nested NER tag

### Big Data & AI Technical Leader

2021-present

2021: VNDIRECT Datawarehouse & AI System

VNDIRECT, Viet Nam

Design and build VNDIRECT data warehouse, develop eKYC system and NLP applications

### Data Science Manager

2018-2020

2020: NewsAI

OSP, Viet Nam

Build a personalized news recommendation product

2019: YottaData Platform

Develop a big data platform for Telcos and Airlines, win Vietnam Digital Awards 2019

2018: OSP Credit Scoring System

Build a credit scoring system leveraging telecom data and social media data

2018: OSP Customer Insight

Construct user profiles for hybrid recommender system

### Solution Leader

2014-2017

2016-2017: Social Box

1link, Viet Nam

Build social monitoring system for mining very large data in social media

2014-2015: Big Crawler

Build large scale and flexible crawler framework using Big Data technologies

## ❖ Publications

**A System for Aspect Based Sentiment Analysis at VLSP 2018** *VLSP'18*  
Vu Anh, Bui Nhat Anh

**Named Entity Recognition in Vietnamese using CRF** *VLSP'16*  
Vu Anh, Le Minh Tuan, Le Minh Hung

## ❖ Certifications

**CCA Spark and Hadoop Developer Exam (CCA175)** *September 2020-present*  
License 100-023-710 cloudera

**Deep Learning Specialization** *February 2018-present*  
License 384YXEJ89KN6 coursera

**Machine Learning** *April 2016-present*  
License MNE2P9TPHPsy coursera

## ❖ Qualifications

*Major languages:* Vietnamese, Python, Java, C/C++, Javascript, R

*Technologies:* Natural language processing, Deep learning, Big Data, Data Warehousing

*Libraries and tools:* Pytorch, TensorFlow, spaCy, pandas, Spark, Hadoop, Hive