

# Anh Tu Nguyen

Ha Dong, Ha Noi, Vietnam  
☎ (+84) 964692796  
✉ [anhtunguyen446@gmail.com](mailto:anhtunguyen446@gmail.com)  
🌐 My Webpage  
🐙 Github   in LinkedIn



## Education

2016–2021: **Bachelor Engineer Degree**, *Posts and Telecommunications Institute of Technology*, Hanoi, Vietnam.

**Major:** Information Technology

Programing, Data structures and algorithms, Analysis & design of information systems

## Research Experience

*Posts and Telecommunications Institute of Technology, Hanoi, Vietnam*

2019–present **Legal Text Processing**, *Research and develop neural network architecture to extract information from legal text.*

- Graduate theis: **Information extraction from legal questions**
- **Develop a bidirectional joint model for intent detection and slot filling**

Advisor : **Dr Ngo Xuan Bach**, *Associate Professor, Vice Dean of Faculty of Information Technology, Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam.* ([Personal Web-page](#))

[VinBrain JSC, Vietnam](#)

2021–present **Medical text processing, audio processing**, *Research, develop and apply state-of-the-art neural network architecture, ML algorithm for medical text processing and audio processing.*

Advisor : **Hoang Vu**, *Head of speech and language processing department, VinBrain JSC, Vietnam* ([Web-page of Company](#))

## Job Experience

10/2020– **Technology Specialist**, *Vingroup Big Data Institute.*

- 1/2021
- Learning the courses Linear Algebra, Probability & Statistics, Machine Learning, Deep Learning, Computer Vision, Natural Language Processing, AI Ethics
  - Work on the Machine Learning course final project 'Vietnamese Text Sentiment Classification'
  - Work on the Deep Learning course final project 'Transferring Images into Vietnamese Traditional Painting Style'

2/2021- **Applied Scientist**, Vinbrain JSC.

- present
- **Automatic Radiology Report Editing Through Voice**
    - Building the NLP controller model based on JointBERT model for detecting intent and extracting content command of doctor
    - Deploy NLP controller and integrate with ASR service
  - **Booking appointment chatbot**
    - Synthesize dataset based on pattern question
    - Building the NLU model to detect intent and extract entity from questions
  - **Stroke detection chatbot**
    - Building neural model to detect stroke through voice record
    - Building machine model use acoustic feature to detect stroke through voice record
  - **Medical equipment registration classification system for ministry of health**
    - Building machine reading comprehension model based on XLM-R model to extract information from document (support both English and Vietnamese)
    - Optimize machine reading comprehension model using knowledge distillation, quantization and transform pre-trained multi language model to bi language model
    - Integrate OCR model, NLP model to classify the power of attorney
  - **Spelling correction service**
    - Crawl and synthesize dataset from online resources
    - Building a joint learning model to detect error and correct error
  - **Automatic speech recognition system**
    - Building a joint learning model to recover capitalization and punctuation
    - Building a spoken norm model to transform spoken language to written language
  - **Search engine**
    - Building sentence representation model based on SimCSE model
    - Integrate sentence representation with FAISS framework to build a semantic search engine

---

## Computer skills

**Programming Languages** Python, JAVA

**Frameworks and libraries** Pytorch, Scikit-learn, Numpy, Pandas, Matplotlib, Flask, Huggingface

**Software Development** Programming Paradigms, GIT, Docker

**Languages** **Native:** Vietnamese, **Fluent:** English

---

## Publications

### In Conference Proceedings

- 2021 **Nguyen Anh Tu**, Hoang Thi Thu Uyen, Tu Minh Phuong, and Ngo Xuan Bach. Analyzing vietnamese legal questions using deep neural networks with biaffine classifiers. In *International Conference on Neural Information Processing*, pages 513–525. Springer, 2021.
- 2021 Manh Hung Nguyen, Vu Hoang, **Tu Anh Nguyen**, and Trung H. Bui. Automatic Radiology Report Editing Through Voice. In *Proc. Interspeech 2021*, pages 4862–4863, 2021.
- 2021 Ta Duc Huy, **Nguyen Anh Tu**, Tran Hoang Vu, Nguyen Phuc Minh, Nguyen Phan, Trung H Bui, and Steven QH Truong. Vimq: A vietnamese medical question dataset for healthcare dialogue system development. In *International Conference on Neural Information Processing*, pages 657–664. Springer, 2021.