

Chin-Ho Lin

📞 (65) 96957432 ✉ chinho.lin@mailbox.org  [chinholin](#)  [tainvecs](#)

Work Experience

- **Data Scientist, Intelllex, Singapore** Jan. 2021 - Present
 - In charge of the search functionality of our platforms, I design and implement algorithms and APIs for search and SEO related features.
 - Collaborate and communicate with designers and legal experts for feature implementation, sprint planning, and quality assurance.
- **Junior Data Scientist, Intelllex, Singapore** Jul. 2019 - Dec. 2020
 - Developed data science microservices for preprocessing, indexing, and ranking user queries or documents.
 - Worked with a DevOps engineer; we handled crawlers, data pipelines, and data access from SQL databases.
 - Created and maintained the Gitlab CI, Dockerfiles, and Kubernetes cronjob and deployment YAML files for our data science git repositories.

Education

- **National Taiwan University** Sep. 2015 - Aug. 2017
 - Master of Science, major in Computer Science
 - Overall GPA: 4.25/4.3, **Rank: 5/131**
 - Thesis: Learning to Map Natural Language Statements into Knowledge Base Representations for Knowledge Base Construction ([LREC 2018](#))

Projects

- **Intelllex: [finreg.sg](#)**
 - A legal knowledge search platform for financial regulation Q&As and documents.
 - Designed and implemented search features, including Elasticsearch query generation, search bar auto-completion for user queries, or suggesting the next articles based on what the user is reading.
 - Implemented features such as summarization or keyword extraction from articles for the Search Engine Optimization of our platform.
 - Developed an online business description analyzer that generates a customized legal Q&A digest based on user input.
 - Keywords: Golang, Elasticsearch, MySQL, RESTful API, Kubernetes, Docker, Cronjob
- **Intelllex: Information Extraction System**
 - Redesigned and implemented the entity extraction module in our query processing pipeline and improved the performance by **22.1%**.
 - Partnered with a legal expert, built a legal case and legislation extractor from free-form documents, and delivered Proof of Concept to our potential clients.
 - Keywords: Python, Text Representation and Similarity, Regular Expression, Kubernetes, Docker, Microservices
- **Side Project: [Tweets Sentiment Classification](#)**
 - Given 200,000 labeled and 1,178,614 non-labeled tweets, I built a binary classification model and tried semi-supervised learning on the non-labeled tweets. The final accuracy reached 0.845.
 - Keywords: Word Embedding, GRU, LSTM, Ensemble, and Semi-supervised Learning

Skills

- Strong: Python, Natural Language Processing, Machine Learning, Elasticsearch, Docker, Golang, Linux, Regex
- Knowledgeable: SQL, Gitlab CI, Clojure, Kubernetes
- Languages: Chinese, English