



Portfolio

西川 荘介

Website <https://sosuke115.github.io/>

Github <https://github.com/Sosuke115>

Twitter https://twitter.com/ponyo_ponyo115

Blog <https://qiita.com/Sosuke115> (Qiita) <https://nchaso.hatenablog.com/> (はてなブログ)

学歴

- ・ 東京大学大学院情報理工学系研究科
電子情報学専攻 4.2020 -
- ・ 東京大学工学部電子情報工学科卒業 3.2020

職歴

- ・ リサーチアシスタント（国立情報学研究所 4.2020 - ）
自然言語処理における敵対的サンプルの研究をしています。
- ・ 機械学習エンジニア（Studio Ousia Inc. 2.2020 - ）
多言語モデルの研究をしています。「森羅2020-ML: 多言語分類タスク」に取り組み、2
言語で優勝しました。<https://nchaso.hatenablog.com/entry/2020/12/12/162906>
- ・ ソフトウェアエンジニア (Stockmark inc. 12.2018 - 12.2019)
ニュース記事をレコメンドしてくれるwebサイトを運営するベンチャー企業での開発経験
があります。利用者が所属している団体がそのwebサイトにおいてどの分野の記事を主に
見ているかのデータを分析し、誰がどの分野に詳しいか一目でわかるようなシステムを作
成しました。webサイトにおいて何を閲覧しているかのデータをmysqlデータベースから取
得し、Pythonにて統計処理を行い、Typescript、HTML、CSSでUI部分を実装しました。
その他簡単なシステム作成やデータ可視化を行っていました。
- ・ 2020年度サマーインターン
機械学習エンジニア (Wantedly Inc. 3週間)
研究開発エンジニア (Cookpad Inc. 2週間)
研究開発エンジニア (Sony 1ヶ月)

研究

- ・ 大学院研究（東京大学、越前研究室 4.2020 - ）
エンティティを用いて多言語処理タスクの性能を向上させる研究を行っていま
す。
- ・ 卒業研究（東京大学、鶴岡研究室 4.2019 - 3.2020 ）
大学4年時は機械学習、特に自然言語処理・ゲームAIを中心に扱う研究室に所
属し教師なし機械翻訳、多言語分散表現を中心に研究をしていました。「教師なし
機械翻訳」の出力結果を分散表現の学習データに用いることで多言語分散表現の質
を向上させる手法を提案しました
https://www.anlp.jp/proceedings/annual_meeting/2020/pdf_dir/P5-20.pdf
<https://arxiv.org/pdf/2006.00262.pdf>

プログラミング

Python・Pytorch・Tensorflow・Chainer（研究開発で日々利用）
Javascript (Typescript)・HTML・CSS・Mysql・AWS（EC2, S3）・GCP (BigQuery)
（インターン、大学の授業で利用）
C・OCaml・Ruby（大学の授業で利用）
Google Apps Scripts（個人にてLINEbot、自動返信サービスの開発）
Atcoder 1003

製作物・達成した事

- ・ 中本分類器 https://github.com/Sosuke115/Nakamoto_Classifier
MLP画像分類器をフルスクラッチで実装
- ・ CLDC <https://github.com/Sosuke115/CLDC>
多言語分散表現評価タスク用にPytorchにて実装
- ・ LINE bot GASを用いてLINEbotを作成
- ・ DQN五目並べAI <https://github.com/Sosuke115/Gomoku-AI>
Chainerにて実装（学生実験）
- ・ 音声で恋愛ゲーム <https://github.com/Sosuke115/dating-simulation>
音声認識エンジン julius、Python、シェルスクリプトにて実装（学生実験）
- ・ ゲームソフト売上数の可視化
<https://www.youtube.com/watch?v=LWtJZZejSDs>
D3.js、HTML、CSSにて実装（学生実験）
- ・ Deeplearning for NLP（松尾研主催、NLP講座）
<https://deeplearning.jp/en/deep-learning-for-nlp-2019/>
Word2vec、LSTM、TransformerなどをPytorchにて実装
- ・ OCR with python（flaskの練習で実装）
https://github.com/Sosuke115/ocr_flask

その他

- ・ 塾講師を二年間やっていました。
- ・ 学部時代アカペラサークルに所属していました。
- ・ IELTS 6.0、オンライン英会話をやっています。
- ・ 学部GPA 3.8/4.2（東京大学教養学部理解一類・工学部電子情報工学科）
- ・ 囲碁アマチュア五段