



# Portfolio

西川 荘介

## 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 - ）  
フェイクニュース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。現在、「森羅2020-ML: 多言語分類タスク」に取り組んでいます。<http://shinra-project.info/shinra2020ml/?lang=en>
- ・ 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。

## 研究

- ・ 大学院研究（東京大学、越前研究室 4.2020 - ）  
深層学習とセキュリティを扱う研究室に所属しています。フェイクニュースを大規模事前  
学習モデル等を用いて検知する手法について研究しています。
- ・ 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。本研究は言語処理学会第26回年次大会で発表しました。またm内容を発展させたも  
のを国際学会COLING2020に投稿予定です。  
「教師なし機械翻訳の出力結果から作成したBilingual Word Embeddingsの分析」  
言語処理学会 第26回年次大会 西川 荘介、鶴岡慶雅  
[https://www.anlp.jp/proceedings/annual\\_meeting/2020/pdf\\_dir/P5-20.pdf](https://www.anlp.jp/proceedings/annual_meeting/2020/pdf_dir/P5-20.pdf)

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

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

**Twitter** [https://twitter.com/ponyo\\_ponyo115](https://twitter.com/ponyo_ponyo115)

## プログラミング

Python・Pytorch・Tensorflow・Chainer（研究開発で日々利用）  
Javascript・HTML・CSS・Mysql（インターン、大学の授業で利用）  
C・OCaml・Cython（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（sosuke115、緑、1003、Python）

## 製作物・達成した事

- ・ 中本分類器 [https://github.com/Sosuke115/Nakamoto\\_Classifier](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
- ・ 言語処理100本ノック（進行中）  
[https://github.com/Sosuke115/nlp\\_100\\_exercise\\_2020](https://github.com/Sosuke115/nlp_100_exercise_2020)

## 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 3.8/4.0（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