Robert **Balayan**

Japan, Ōsaka-fu, Ōsaka-shi, Jotō-ku Furuichi 1-21-22 #102
+81 80 7802 6118 | robert@balayanr.com | balayanr

AI and Software Engineer in Osaka, Japan. Looking to work with Data and Machine Learning. Interested in Database Architecture and Distributed Systems.

summary

- Experienced in Machine Learning, Natural Language Processing and Computational Linguistics, General AI Theory, and Knowledge Representation.
- Programming Languages: Python, SQL, PL/pgSQL, C, Java, Matlab, R.
- Natural Languages: English, Russian, Armenian (Conversational), Japanese (Basic).
- Frameworks and tools: NumPy, TensorFlow, keras, AWS (RDS, EC2, DyDB, S3), Luigi, Selenium, BeautifulSoup, Kubernetes, Docker, Hadoop, NLTK, Django REST, LaTeX, Draw.io for ERD design.

work experience

April 2018 - February 2020

LEAD AI ENGINEER
Kuon Technologies, Osaka, Japan (クオンテクノロジーズ株式会社)
Lead engineer of the AI of the company's main product - Beep Shift, a scheduling AI that assigns part-time employees to the positions required by different businesses. The product is in its final stages of development with further expansion planned to include sales and requirement prediction.
Handled everything related to the AI side of the project: from database architecture and API design to AI development and deployment.

February 2020 - Now

LEAD SOFTWARE ENGINEER
K.K. Tree, Osaka, Japan (株式会社ＴＲＥＥ)
Lead the team handling web-scraping. Designed the whole ETL pipeline around the Luigi framework, which allowed us to maximize server utilization with great scalability as most tasks can run in parallel.
Designed and implemented a Python framework that allowed for new crawlers to be created in minutes with less than 20 LOC while still providing great flexibility for complex websites.
Designed the Database used for the project in PostgreSQL, with S3 for file storage.

Projects are explained in greater detail in a section below.

education

2014-2017

HONOURS BACHELOR OF SCIENCE
University of Toronto, Canada
Computer Science Specialist with Focus in Artificial Intelligence
Fields studied:

- Machine Learning, Convolutional and Recurrent Neural Networks
- Computational Linguistics and Formal languages, Natural Language Processing, Phonetics
- General AI theory, Knowledge Representation and Reasoning
- Graph Theory and Algorithm Design
- Statistics, Linear Algebra and Multivariable Calculus

Relevant Projects:

- CSC412 Solo Research Project: Exploiting Structure For Classification Of Handwritten Japanese Characters
- CSC384 Group Research Project: CSP Algorithm Analysis
- CSC343 Solo Project: SQL-based course recommendation engine

2017-

CERTIFICATIONS
Coursera
Hadoop Platform and Application Framework, License D6W6PEBPALWW

skills and experiences

- 6+ years of experience **programming in Python**: from manipulating SQL databases to Neural Networks and Deep Learning using Numpy, TensorFlow and Keras to Chart parsing and word sense disambiguation using NLTK to Data Mining with Selenium and BeautifulSoup. This is my language of choice for both personal projects and research.
- Interest in Natural Language Processing and Computational linguistics. Familiarity with basics of Information Retrieval, Speech Recognition and Synthesis, Word sense disambiguation, statistical chart parsing disambiguation, and Question Answering.
- Experience designing relational databases and implementing them in PostgreSQL, MySQL and SQLite. Experience with AWS storage solutions like DynamoDB and S3. Experience with **HDFS and HBase** as part of Hadoop Platform and Application Framework certification.

projects

February 2020 - Now

Web scraping @ K.K. Tree

Designed and implemented an ETL pipeline for a website that unifies listings from government tender websites. The project is designed around **Luigi**, which schedules tasks to maximize server utilization. Each task handled a different website: loaded and navigated the page with **Selenium**, parsed the data by locating elements with XPATHs and used **Tabula** and **PDFMiner** to extract data from PDFs. The data is then handed back to Luigi in CSV files to be inserted into the **PostgreSQL database** in batches to optimize load. **Psycopg2** is used to get data from the database, **Boto3** is used to talk to AWS S3 to store PDFs and screenshots.
Deployed on AWS EC2; Database is in PostgreSQL, deployed on AWS RDS; AWS S3 is used to store files for listings.

November 2019

Crawling of Hello Work's Job Listings.

A customer who wanted a page that only contained listings relevant to their services.
Everything was deployed on AWS Lambda, with functions handling both server-side rendering for pages and crawling. Used BeautifulSoup for page parsing and Selenium to help with navigation, database in PostgreSQL and deployed on AWS.

June 2018 - January 2020

Beep Shift @ Kuon Technologies

At Kuon Technologies I worked on creating an AI that would assign employees to shifts created by the manager. The shifts were represented as a step curve, with number of people required at each time. The AI would transform that into a Constraint Satisfaction Problem, with time constraints on employees (Overtime, maximum salary, etc) and on the shop (budget, position requirements). The model also allowed for managerial preferences, which the AI interpreted as a selection order via a heuristic function. Once the AI was ready, our customers provided us with text files with raw receipt printout data, which I had to parse and load into a database, perform feature selection and extraction from the parsed data and create a model that would predict the number of employees required to run the shop.

CSC412

Solo Research Project: Exploiting Structure For Classification Of Handwritten Japanese Characters.

- Developed a parser for the Electrotechnical Laboratory datasets, which come incredibly compressed and required a bespoke decoding library.
- Tested 4 different architectures with 2 different goals: direct kanji classification and multilabel classification of radicals that make up the kanji.
- Used AWS EC2 p2.xlarge and c4.8xlarge instances to train the models in parallel.
- Compiled the report in LaTeX, generally following standards used for publications.
- Published all source code and the report on Github

CSC384

Group Research Project: Constraint Satisfaction Problem Algorithm Analysis.

Was the *de facto* team leader: provided the team with a selection of algorithms outside of the course's scope that would be interesting to analyze. Final selection was iterative improvement, path consistency, and a tree-specific algorithms.

languages

English	Fluent
Russian	Fluent
Armenian	Conversational
Japanese	Basic

interests and hobbies

Artificial Intelligence: I have always been interested with AI and inspired by its portrayal in science fiction. It is my dream to work on implementing a general AI. One of the most interesting roadblocks to me is the way humans interact with AI: in science fiction it is almost always seamless, like talking to another person. Personally I have always had interest in Natural Languages, so this is one of the fields I would like to study in depth.
Sports: I am an avid cyclist, runner and hiker. I used to train with the University of Toronto Road Racing team and participate in various local cycling events. Hiked to the top of Mount Fuji, as well as many mountains in Kansai region. I also enjoy diving and have various advanced diving certifications.
Traveling: I enjoy traveling whenever I have the opportunity. I have visited many countries, including Japan, Spain, France, Italy, England, Germany, Egypt, Turkey, Georgia and Armenia. Lately, I have been traveling with my road bike and touring in the places I visit.
GunPla: One of my favourite hobbies is building plastic model kits. Over time I have built dozens of kits from Mobile Suit Gundam and Warhammer 40K figurines.