

KASIM CHRISTIAN LOAN

COMPUTER SCIENCE MAJOR

christian.kasim.loan@gmail.com

Website <https://ck1-it.de>

Mobile : 01784377548

ABOUT ME

Language Skills Scala, Java, Python, SQL, Bash, C/C++, HTML/CSS/JS, Haskell, Prologue, JSP, PHP, Liquid, Ruby

Big Data Skills Spark, Kafka, Hadoop, Kubernetes, Ambari, Hortonworks, Databricks

ML Skills Skills Tensorflow, Keras, Pandas, Sk-learn, Seaborn, Deeplearning4j, Numpy

DevOps Skills Kubernetes, Docker, Azure, Powershell, ARM, awk, AWS, Jenkinx-X, Grafana, Prometheus, VMs

Web Skills SpringBoot, Nodejs, Jekyll Django, Bootstrap, Jquery, Shopify, Woocommerce, Wordpress, Front Matter

Personal Interests Coding, Reading and writing sci-fi, meditation, yoga, martial arts

WORK EXPERIENCE

Daimler Center for Automotive IT Innovations

Big Data/ Cloud Engineer and ML Engineering

Berlin, Germany (Febuary 1. 2019 - present)

- Currently architecting and implementing a big Data ML pipeline and analytics application for streamed car IOT data for Daimlers automotive driving fleet, with Spark,Kafka and Hadoop inside of Azure with Kubernetes and Docker
- Managed the deployment of software packages with Docker, Vm's, Azure and ARM files
- Predicted car lane changes on basis of drone surveillance data of German highways.
- Deployed Infrastructure as code in Azure with ARM files
- Deploying Machine Learning Models on structured streamed data sets in Spark with Scala, Python and Kafka inside of Azure in a Kubernetes cluster. Using Docker to define containers and dependencies
- Building a next generation Neural Network, "Spatio - Temporal Graph Convolutional Neural Network", which is a combination of LSTM + CNN, which takes in flow graphs as input. Using it, to predict traffic flow world wide in real time.
- Wrote custom distributed Kubernetes Framework
- Defined a Prometheus and Grafana Monitoring and Metrics system which ingested data from Kubernetes cluster components like Kafka and Spark.

GT-ARC and Distributed Artificial Intelligence - Lab / DAI - Lab

Big Data and ML Engineering

Berlin, Germany (Aug. 2017 - December, 2018)

- Build the backend foundation for a distributed meta machine learning streaming platform, "CODA." It is a tool that learns which ML problems are solved with which algorithms and hyperparameters.
- Set up Spark, Kafka and Hadoop big data cluster infrastructure and integreated them with each other. Also provided handy usage documentation
- Wrote custom serializers for the data streamed from Kafka topics and made them easily available for cluster computing
- Set up Hadoop on our cluster and finished building our foundation for the ML dataset repository after integrating Kafka into Spark
- Used Spark to ingest and feature engineer datasets larger than 1TB into Hadoop cluster
- Meta machine learning experiments with python
- Created the frontend and backend for an emotion recognizing face scanner app which uses a webcam. You ask the app to tell you a joke. It tells you a joke and scans your face to detect emotions. Using the recorded emotions, the app creates a database of the best jokes

Moebel 24 - SEO Startup

Backend, Automated Testing

Berlin, Germany (Nov. 2016 - Mar. 2017)

- Java frontend and backend test development with Selenium
- CMS Administrating, JSP development, SEO optimizations, System and network administrating

FREELANCE WORK EXPERIENCE

Spark NLP / John Snow Labs

Scala Spark ML Engineer

Remote, Berlin, Germany (2020 Januar - Present)

- Extending one of the most used AI Frameworks-Libraries frameworks in the enterprise domain
- Defining Scala Spark pipelines for NLP models
- Exporting Tensorflow models to Spark Scala, distributing and optimizing them like Elmo, Albert, XLNet and more

Avangrid

Data Science consulting and Scala Spark ML Engineer

Remote, Berlin, Germany (2019 Dec - Present)

- Analyzed time series IOT database to predict for death of machines(aka machine deterioration) a customer and architected a Scala Spark Machine learning pipeline for it
- Implemented the scala spark ML pipeline and deployed the machine deterioration prediction Model into their system
- Deployment of Jar application to Ambari
- Currently providing fruther support and guidance for the company

HACKATHONS

Microsoft x Redbull - Enhance Student life with Azure - 2019

- Won hackathon with a "Sentiment lecture analysis with demographic predictions" pipeline
- Architected the pipeline, implemented it in Azure and trained Tree Model to predict demographics
- The presentation is visible here <https://youtu.be/kFMga8kdl2U>

Rohde Schwarz Engineering Competition 5G Broadcast - 2019

- Placed as best team in a global competition
- Predicted Time Series data, similar to viewing figures of multiple TV channels for 3 months into the future

EDUCATION

Coursera - Deep learning specialization

Coursera (2018)

- Nano-Degree Neural Networks and Deep learning
- Nano-Degree Improving Deep Neural Networks: Hyperparameter tuning, Regularization and Optimization
- Nano-Degree Structuring Machine Learning

Technische Universitaet Berlin

Berlin, Germany (2014 - present)

- Computer science major, emphasis on AI/MV/ML

Humboldt Universitaet Berlin

Berlin, Germany (2013 - 2014)

- Two semesters as computer science major

Theodor-Heuss Gymnasium

Berlin, Germany (2010 - 2013)

- Abitur

NLU/NLP

- Used Rasa NLP to extract and learn entities and intents for a internet support and phone consultation for Deutsche Telekom
- NLU used Rasa Core to define actions based on extracted intents and entities like suggesting a phone or giving a solution to a router problem
- The chatbot takes in a message and uses a knowledge graph to find phones that match the users criteria
- Build a real time topic segmentation video based on a lecturers voice input. First the audio is translated to text and then the text will be segmented by topic

Machine Vision

- Build custom CNN in Keras from scratch to classify food from 4 different classes with accuracy 80%
- Used a pre-trained CNN with open CV to build a cup-recognizing-kettle which automatically boils water if you set a cup in front of its camera
- Implemented style transfer with Keras
- Used the Stanford Affectiva face emotion scanner API to build a emotional joke rating system web app.
- Build a "Presentation engagement analysis tool" App for a Microsoft Hackathon and won. It analyzes a person presenting to an audience. You can find it here <https://youtu.be/kFMga8kd12U>

Time Series Analysis

- Led a team of 5 students to implement a state of the art paper for time series analysis
- Build a streaming time series analysis ML-engine for electrical appliance data. Used an ensemble of KNN algorithms and boosting mechanism to achieve 98% accuracy for classifying electrical devices based on their electrical fingerprint.
- Implemented credit card fraud detection with 99% accuracy and about 80% F1 accuracy
- Predicted 5G stream usage and optimal advertisement placement for 3 months in the future for a global data science hackathon, won first place.
- Building the streetnetwork as a graph and using a time serieses of graphs to predict traffic flow with a Spatio Temporal Convolutional Neural Network

RECENT PERSONAL PROJECTS

Time series machine learning engine

(Summer 2018)

- Designed software architecture in UML for state of the art time series machine learning engine based on the scientific paper released 2018 by IEEE members
- Lead a team of 4 students, segmented the paper into work packages for fellow students and managed them
- Wrote all of the machine learning code and most of the database code which communicate with a smart power outlet that streams electrical consumption data over wi-fi
- Extracted unique time series fingerprint patterns from the smart power outlet and used them to build a machine learning algorithm that can correctly classify any appliances based on this architecture
- Entire system is designed for online learning, making the model perform better over time

Telekom NLP/NLU chatbot

(Summer 2018)

- Tasked to design a chatbot for Telekom which would interact with people who call tech support; also helps users choose which cell phone is best for them
- Used a dataset of around 10 million forum posts from the Telekom tech support forum and extracted problem classes and correlated solution classes
- Implemented it with Rasa NLU/NLP framework which enabled us to quickly use state of the art algorithms

Machine vision - Automated mug-recognizing kettle

(Summer 2018)

- Set up a headless Raspberry Pi with a CNN that is fed a camera stream from a cam that is also connected to the Raspberry
- A radio frequency transmitter is attached to the Raspberry, making it able to communicate with a smart power outlet into which a kettle is plugged
- When the CNN recognizes a mug, it sends a signal with its transmitter to the smart power outlet that is connected to the kettle, making the water boil

Fullstack Development - Precise model of human joint and bone system

(Summer 2018)

- Created the prototype for a highly precise. simulation of the human bone and joint system in three.js
- Joints and bones are rotatable and in parallel the center of masses for every body part is calculated
- Used to mathematically and visually prove Aikido concepts for a book of a rheumatology doctor and math master.

Chat bots

(Summer 2017)

- Created a bot for telekom, which can handle customer support and questions about phone specs
- Wrote chat and interaction bots for a few social media and dating platforms

Webcrawlers

(Winter 2017)

- Built a webcrawler that sucks all YouTube videos from an input subreddit list and uploads them to a YouTube channel at scheduled intervals
- Created a webcrawler that crawls every of Amazons categories (over 10k) and for every of the categories it crawls the products in the top100 of that category. Based on the top100 categories, I calculated meta data based on the reviews and scores to identify potentially low competition and high profit categories. The crawled data is rendered to a frontend via React, currently working on this to commercialize it.
- Analog project for ebay-kleinanzeigen.

E-Commerce shops

(Winter 2017)

- Built a watch and an anime e-commerce store from scratch with Shopify using custom written liquid code
- Successfully used Facebook ads to generate profitable traffic by using Pixel tracker techniques
- Created the backend for a crypto voucher store which was connected to multiple crypto exchanges and allowed users buy vouchers for crypto coins which they can later redeem

AWS GAN

(Winter 2017)

- Started building A GAN in AWS using Keras and Python.
- Generated Images of Food

TECHNOLOGICAL EXPERTISE

Scala

- Implemented multiple streaming ML data pipeline with scala from Kafka input with watermarks and aggregations
- Batch Feature Engineering with Spark
- Deployed a Temporal-LSTM from Keras on a real time Kafka stream
- Writign library code for Spark NLP for lots of unstructured data
- Many spark ETL jobs

Java

- Deployed and documented Spark/Kafka/Hadoop/ cluster enviroment Java for multiple laboratories
- Wrote Kubernetes distribution Framework
- Wrote many Spark ET jobs
- Custom serializers for Kafka and Spark
- Created handy functionality wrappers for Spark/Kafka/Hadoop
- Feature engineering with Spark
- Code deployment in clusters with Maven
- Implemented at that time state of the art ML time series algorithm for streaming electrical data based on from IEEE members. The paper can be found here <https://ieeexplore.ieee.org/document/8118142/>
- Built stock market simulation with JMS and ActiveMQ
- Built a few simple Android apps and a game

Python

- Tensorflow, Keras, Numpy, Scikitlearn, Pandas, Matplotlib ML tool kit
- Feature Engineering and Model deployment with Pyspark and many ETL jobs
- Built a GAN with keras for generating food images
- Django websites
- Content crawlers and uploaders for many e-commerce websites
- Chatbot for Deutsche Telekom for their customers. The chatbot worked on 2 domains, internet tech support and phone buying consultation. Implemented with Rasa NLU/Core
- Messaging bots for Instagramm and Facebook
- Created simulations for ants colonies
- Build simulation for auctions, distributed interacting agents
- Building a content crawler for reddit and recombining the content to build niche websites and Youtube channels in an automated fashion.

C/C++

- Implemented data structures like trees, AVL trees, ques, and many search algorithms
- Wrote servers like DHCP and simple email servers
- Loaded a trained CNN with Open CV and used it to classify live streamed objects for the automated MV water cooker
- Participated in the development of Cube Universe Game : <https://cubeuniverse.net/>

Javascript/Node js / React

- Built Restful servers, Email servers, User Authentication
- Endpoints for frontend to query DB
- Build a highly precisce simulation of the human joint and bone system which is manipulable through a GUI, made with three.js
- Build Social media clones
- Crawled metadata from e-commerce websites and renderd the data to a React frontend.

Databases

- Built and administrated distributed databases with Hadoop
- Build and deployed DBs with Hadoop, Mongo , Derby, SqlLite, SQL etc ..
- Created restful endpoints to query databases for frontends

HTML/CSS

- Build the fullstack single page application for the emotional Joke rater machine
- Responsive websites with Bootstrap

Liquid

- Created a watch and an anime merchandise store and tweaked themes for them

PERSONAL DATA SCIENCE PROJECTS EXPERTISE

- Image Food class detection. Using Simple CNN
- Credit card Fraud detection. Using SVM
- Time Series IOT electrical data Classifier. Using Ensemble KNN Method.
- Chatbots for customer service and product consultation. Using RASA LSTM Chatbot framework.
- Handwritten digit classification. Using simple MLP/CNN
- Facial Sentiment Analysis. Using Affectiva Emotion API
- Twitter Text Sentiment Analysis. Using LSTM
- Video topic segmentation and audience sentiment analysis. Using Azure ML Studio
- 5g time series Internet data usage prediction. Using ARIMA and Fourier Series.
- Car time series behaviour prediction from Drone Bird's eye highway dataset, using XGboosted trees and Grid-search Methods.
- "Spatio - Temporal Graph Convolutional Neural Network", LSTM+CNN to predict flow network graph features.