RAESETJE BONJO SEFALA

raesetje.sefala@gmail.com
https://github.com/sefalab/gis_x_vision

WORK EXPERIENCE

Aug 2021-Present

Research Fellow, Distributed AI Research Institute |PI - Dr. Timnit Gebru

- Building and publishing a dataset of annotated satellite images for the purpose of analyzing the effects of spatial apartheid in South Africa.
- Building Machine learning models to detect and classify neighborhoods on satellite images according to their types(suburbs, townships, informal settlements, e.t.c.)
- Design and implement software and databases to efficiently build models and evaluation pipelines for large image datasets(over 10 million images)
- Analyzing the effects of spatial segregation over a period of time on satellite images using mainly using computer vision techniques and techniques from other disciplines.

Oct 2020-Jun 2021

Research Intern, Mila- Quebec AI Institute | PI - Prof. Yoshua Bengio

- Our goal is to create data-oriented learning algorithms to help us create more representative poverty estimation maps in the urban parts of Nigeria using satellite images and other spatial datasets.
- **Role**: Compiling ground truth datasets (from different sources such as household surveys, and geospatial data layers), and then building data-oriented models to estimate poverty in the urban parts of Nigeria using satellite images.
- **Skills:** Curating spatial datasets for machine learning use, Modifying convolutional neural networks so that they incorporate other spatially related attributes for improved performance; Using image clustering techniques for creating data set splits to efficiently create maps for different landscapes; Building data-oriented models; Optimized the models using various parallelization and vectorization techniques to make the code run faster and more efficiently.

Jun 2020-Sep 2020

Data Scientist, Data-Intensive Development Lab- UC Berkeley |PI- Prof. Joshua Blumenstock

- Using Machine Learning and spatial datasets to build poverty estimation maps for the Nigerian Government to assist them to identify the poorest regions in Nigeria so that they can prioritize them for their COVID19 relief grants.
- **Role**: Finding creative ways to create labels that are representative of the current ground truth, train machine learning models and evaluate the results, communicate the results, get feedback and then improve the models/datasets accordingly.
- **Skills**: Data cleaning and manipulation using python packages; Used machine learning techniques to extract image features, do transfer learning and modified state of the art models to work on particular datasets; Built Regression and classification models using various data types to make poverty estimates in different social classes; Created reports and visualizations of the data and findings; Explored various data splitting methods for the big datasets we were dealing with(+1 million images); Optimized the models using various parallelization and vectorization techniques to make the code run faster and more efficiently; optimizing code to run efficiently across multiple CPUs and GPUs.

Aug 2019-
Jan 2020

Data Engineer, Targeting Talent Programme (TTP)- Wits University

- Created a database that tracks all of the program's alumni since 2009, this is used to measure the program's impact. Created a data version control system and also integrated a data analysis tool.
- **Role**: Consolidating multiple surveys from excel sheets into a single database; Created a pipeline for data capturing.
- **Skills**: Efficient data cleaning and manipulation using python; Creating a data capturing interface using Python and PostgreSQL; Version control, Creating Reports and Visualizations using Tableau.

May 2018-
Aug 2018

Data Scientist, Data Science for Social Good fellowship, University of Chicago

- The Data Science for Social Good Fellowship is a University of Chicago summer program to train aspiring data scientists to work on data mining, machine learning, big data, and data science projects with social impact. Working closely with governments and nonprofits, fellows take on real-world problems in education, health, energy, public safety, transportation, economic development, international development, and more. (https://dssg.uchicago.edu/)

- **Role:** We used object detection and classification to find objects, semantic segmentation to determine where they are (on-road/ sidewalk/ wrong lane) and optical flow to determine movement to build a video-processing pipeline for the city of Jakarta to extract structured information from raw traffic footage. (https://github.com/dssg/jakarta_smart_city_traffic_safety_public)
- **Skills:** Liaising with project partners to gather project requirements; Setting up cloud services for a machine learning project; Designing and developing efficient data storage pipelines; Video data cleaning and manipulation; Designing and developing a data labeling pipeline for multiple labelers; Adapting State-Of-The-Art models to solve specific problems; Designing model evaluation pipelines; Writing an academic paper.

Feb 2018-
May 2018

Machine Learning Developer, Wunderman SA

- Part-time work. Built a computer vision application to classify pictures of dog breeds and predict constituent breeds making up mixed breed dogs. In this work, I used Keras and Tensorflow to build a CNN model with GAP layers to perform object localization and then blurred parts of the image to get constituent breeds making up the mixed-breed dog.

Feb 2018-
April 2018

C# Developer, Water reuse sustainability assessment tool

- Part-time work. Built an end-to-end system for assessing water reuse sustainability based on civil engineering Ph.D. candidate Abiola Abimbade's model using C#.

Jun 2017-
Jul 2017
And

Nov 2017-
Feb 2018

Data Scientist, Data Science for Impact and Decision Enablement (DSIDE), Council for Scientific and Industrial Research (CSIR)

- Data Science for Impact and Decision Enablement (DSIDE) is a vacation work program hosted at the CSIR to support capacity building in the ever-growing field of data science by scheduling recruits to participate in mentor-guided and learn-by-doing problem solving of real-world needs as presented by different stakeholders including municipalities, government departments, energy, academics and more. (http://dsideweb.github.io/)

- **Role:** We used an active learning approach to build an image interest ranking system. Some of the tools we used are a Bayesian ranking algorithm to give scores and precision, a CNN for feature extraction, and a Gaussian Process model to smooth the scores taking the features into account. The tool ranked images according to a domain expert's subjective interest and also highlighted the specific content making these images interesting.

May 2016-Jan 2017 **Developer, Blue Ocean VR**

- Help build a strategy to promote the new company
- Built a treasure hunt augmented reality app for Blue Ocean VR to promote their Virtual reality business at the rAge gaming expo 2016

May 2016-Jan 2017 **Robotics interest group host, Mathematical Sciences Support, University of the Witwatersrand**

- We used inverted learning to teach an introductory course in robotics to a general audience of students from different backgrounds. (https://www.youtube.com/watch?v=_cZUwe0Id1Q&t=17s)

Aug 2015-Oct 2015 **Web Development interest group host, Mathematical Sciences Support, University of the Witwatersrand**

- Teaching an introductory course in web development to a general audience of students from different backgrounds.

Feb 2015-Apr 2016 **Lab technical assistant, Mathematical Sciences Support, University of the Witwatersrand**

- Software installations on Mathematical Sciences laboratory computers
- Fixing Mathematical Sciences laboratory computers
- Maintaining research production servers
- Tutoring for conferences (e.g. tutoring for big data conferences, we assist attendees with using R, PostgreSQL, and Hadoop)

Jun 2014-Jul 2014 and Dec 2014-Jan 2015 **Jnr Developer, Vacation work, Sentech**

- Monitor and assist with research and upgrades of business applications
- Shadowing on managing the company's backup and storage systems
- Perform Network troubleshooting with configurations
- Learn how business processes are carried out and how the company implements

EDUCATION

Feb 2018-Feb 2021 **MSc Computer Science**, *University of the Witwatersrand*, Johannesburg

- **Topic**: Using satellite images and computer vision to study the evolution and effects of spatial apartheid in South Africa.
- **Advisors**: Dr Richard Klein, Nyalleng Moorosi, and Dr Timnit Gebru
  This project explores South Africa using satellite images. One of the main problems South Africa is grappling with is how to remove many of the legacies of Apartheid - a former policy of political and economic discrimination and segregation against non-European groups in South Africa. For example, aerial photographs taken by photographer Johnny Miller show the legacy of spatial apartheid - completely

segregated communities of townships next to gated wealthy neighborhoods that have largely remained unaffected by the ending of apartheid [https://www.citylab.com/equity/2016/06/apartheids-urban-legacy-instriking-aerial-photographs-south-africa-cities-architectureracism/487808/]. Our research uses spatial data and machine learning to analyze satellite images of South Africa from 2006 to 2017.

- Together with the satellite images, we used census and connectivity data to curate a country-wide ground truth dataset classifying all the neighborhoods in South Africa according to their types, it shows you where all the townships, suburbs, informal settlements, villages, e.t.c are to enable further analysis of attributes such as resource allocation per neighborhood type(e.g. schools, hospitals, e.t.c.).

Feb 2017-Dec 2017 **BSc (Honours) Big Data Analytics, University of the Witwatersrand, Johannesburg**

- **Relevant Modules:** • Computer Vision •Machine Learning • Data Analysis and Exploration • Distributed Computing • Discrete Optimisation • Data Visualisation and Communication •Introduction to research methods • Research report for Big Data Analytics
- **Final year project:** Investigating different CNN architectures on the task of action recognition in videos (https://github.com/sefalab/Honours-Research-Project)

Feb 2014-Dec 2016 **BSc Computer Science and Information Systems, University of the Witwatersrand, Johannesburg**

- Awards: 9 Certificates of first-class in several courses (6 in CS, 2 in Maths, 1 in Electrical circuits)

PUBLICATIONS & PREPRINTS

- **Sefala, R.**, Gebru, T., Moorosi, N., & Klein, R. (2021, August). Constructing a Visual Dataset to Study the Effects of Spatial Apartheid in South Africa. In Thirty-fifth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (Round 2). 2021
- Nsoesie, E. O., Sy, K. T. L., Oladeji, O., **Sefala, R.**, & Nichols, B. E. (2020). Nowcasting and forecasting provincial-level SARS-CoV-2 case positivity using google search data in South Africa. medRxiv. **(Under Review for journal)** 2020
- Caldeira, J., Fout, A., Kesari, A., **Sefala, R.**, Walsh, J., Dupre, K., ... & Imtiyazi, M. A. (2019, September). Improving Traffic Safety Through Video Analysis in Jakarta, Indonesia. In Proceedings of SAI Intelligent Systems Conference (pp. 642-649). Springer, Cham. 2019
- Burke, M., Mbonambi, S., Molala, P., & **Sefala, R.** (2017). Rapid Probabilistic Interest Learning from Domain-Specific Pairwise Image Comparisons. arXiv preprint arXiv:1706.05850. 2017

ACHIEVEMENTS

- Recipient of the Most Rigorous Researcher Award by the South African Deep Learning Indaba X Community — 2021
- Keynote speaker at the Computer Vision for Global Challenges Workshop at CVPR. LA, USA — 2019
- Oral presentation at the AI for Social Good Workshop at NeurIPS — 2018
    - Won Highlighted paper award at AI for Social Good workshop
- Oral presentation at the 2nd Black in AI Workshop at NeurIPS. Montreal, Canada — 2018
- Recipient of the best poster presentation prize at the Deep Learning Indaba — 2018
- Data Science for Social Good Fellow, at the University of Chicago — 2018
- Recipient of the Sasol Inzalo Foundation scholarship — 2014 - 2017
- Received 9 certificates of first-class from Wits university — 2014 - 2016
- Member of the Golden Key International Society — From 2014

COMMUNITY SERVICE

Co-founder: Women In Computational Science Research (WiCSR) — 2019- Present

- WiCSR is a community that empowers and encourages women's growth and participation in the field of Computational Science Research. This community is for minorities within the field of computational science research which includes but is not limited to: computer science, applied mathematics, computational mathematics/statistics, machine learning, data science/mining, robotics, AI, and any other related field.

Mentor for several undergrad students, University of the Witwatersrand — 2018-2020

- I mentor undergraduate students on both personal projects and school projects which involve computer vision, machine learning, or general software development.

PROGRAMMING LANGUAGES AND FRAMEWORKS

- Currently proficient in Python, Keras, OpenCV, Tensorflow, Dash, PostgreSQL, C#, Java, HTML, CSS, MySQL, GCP, Linux, Tableau, QGIS, GDAL, Rasterio, and Php.
- Have worked with C, C++, and JavaScript.

REFERENCE

- References available upon request.