William Lee

St. Anne's College, Oxford, UK
cs.swarthmore.edu/~wlee1 ◇ wlee1@swarthmore.edu

EDUCATION

Palo Alto High School *2016 (3.97/4.00)*

St. Anne's College, Oxford (Maths, Computer Science) *2018 – Present*

Swarthmore College (Math, Computer Science) *2020 expected (3.9+)*

WORK EXPERIENCE

Search Relevance, Salesforce *Summer 2018*

— Redesigned the search metrics pipeline for Salesforce Search Cloud using Splunk and Hadoop.
— Data analysis on terabyte-scale logs to accurately measure Salesforce Search's performance.

Laboratory for Advanced Sensing, NASA Ames *Summer 2017*

— Designed and implemented a computer vision alignment pipeline to fully georectify 4k UAV footage onto satellite imagery.
— Implemented a sliding window CNN using Keras to classify coral reef morphology (baseline model).

Intelligent Robotics Group, NASA Ames *Summer 2016*

— Developed a entropy-based adaptive gaussian blur module for NASA's stereogammetry suite.
— Full stack web development on NASA's GeoCam space project and xGDS mission planning software.

PROJECTS

Weight Uncertainty in Neural Networks
Reimplemented and replicated the results found in *Weight Uncertainty in Neural Networks (Blundell et al., 2015)*. Extended results by using modeled uncertainty in an active learning framework.

Capturing Population Events Using HMMs
Analyzed genomic sequence data from different human populations in the 1000 Genomes Project. Combined multiple projects from a semester-long Bioinformatics class to create an end-to-end genomic pipeline that converts raw sequence data to population-size estimates using TMRCA. Successfully captured the out-of-Africa bottleneck; concluded that smaller-scale population events require more data.

Synthetic Datasets Using Deep Generative Models
Explored training neural networks on Synthetic Datasets created by Deep Generative Models (GANs, Deep Boltzmann Machines) for my final project in machine learning, with two group members. Created synthetic datasets for MNIST and CIFAR-10 datasets and analyzed efficacy of models trained on generated datasets.

COURSEWORK

University of Oxford	Advanced Machine Learning, Probability & Computing, Probability, Artificial Intelligence
Swarthmore College	Machine Learning, Computer Vision, Bioinformatics

SKILLS

Python, Pytorch/TensorFlow, OpenCV, Sklearn, Java/C++/C, Hadoop/Splunk, Photography, Math