John-Alexander Hall

Final Capstone Proposal

Cervical Cancer Screening

Cervical cancer is a high-risk disease for many women, especially for those in rural areas. Fortunately, in its pre-cancerous stages, cervical cancer is easy to prevent. However, determining the correct method of treatment is challenging due to physiological differences amongst patients. The position of a woman’s cervix greatly influences the type of treatment that proves most effective; a treatment that works for one patient may increase the health risks for another. The dataset provided contains images for three different types of cervixes, all are considered non-cancerous/normal. However, the transformation zones of these cervixes are not always visible, requiring further tests. The purpose of this capstone is to create an algorithm that correctly identifies the transformation zones to aid healthcare practitioners at screening.

The dataset was provided by MobileODT, a healthcare company that creates medical devices used for patient assessment and visualization. There are several gigabytes of image data within the zipped files that can be used to build the algorithm. I can access the images using the wealth of image reading modules found in the python data science ecosystem.

The most prominent technique that I plan to use from this course during this project is deep learning classification. Multi-layer Perceptrons, Convolutional Neural Networks, and Recurrent Neural Networks are the best image classifiers I have implemented and I plan on making full use of them.

The biggest challenge I plan on facing is computational. Working with image data and neural networks requires a lot of computing power and time. My local machine may not perform as well I hope, forcing me to explore other avenues like Amazons EC2 resources.