Malaria Detection with Machine Learning

Improving the worlds malaria detection systems through artificial intelligence and machine learning.

Executive Summary Problems

Malaria

- Malaria is caused by **Plasmodium parasites** which are transmitted through mosquitoes.
- Despite our best efforts there are more than 229 million cases yearly with over 400,000 deaths.
- Two third of malaria deaths are children under the age of 5.

Detection

- Normal malaria diagnostics require careful inspection from an experienced lab professional.
- This process is extremely time and resource intensive, as it requires many personnel and expensive lab equipment.
- The accuracy is completely dependent on human expertise.

Executive Summary Solutions

Findings and Takeaways

- Machine Learning systems can properly handle and sort patient data.
- Artificial Neural Networks can accurately and consistently diagnose patient data into ‘Uninfected’, and ‘Parasitized’ groups.
- Machine learning models can be replicated and deployed globally to combat the world malaria endemic

Next Steps

- Improve the model by having additional professionals review and test the model.
- Apply the model on numerous data sources to test its global efficacy.
- Assemble both technical and regulatory teams to build the most accurate model possible, and have the model recognized and approved by regulatory agencies.

Problems and Solutions

Problems

- Current diagnostic techniques are tedious and expensive.
- Drain personnel and resources from working on treatment.
- Human error can lead to mistakes.

Solutions

- Implement an Artificial Neural Network to practice sorting patient data between ‘uninfected’ and ‘infected’ images.
- Use Convolutional Neural Networks to maximize the accuracy and recall capabilities of the model.
- Test the model multiple times to improve the model’s ability to produce the most accurate results.

Solution Viability

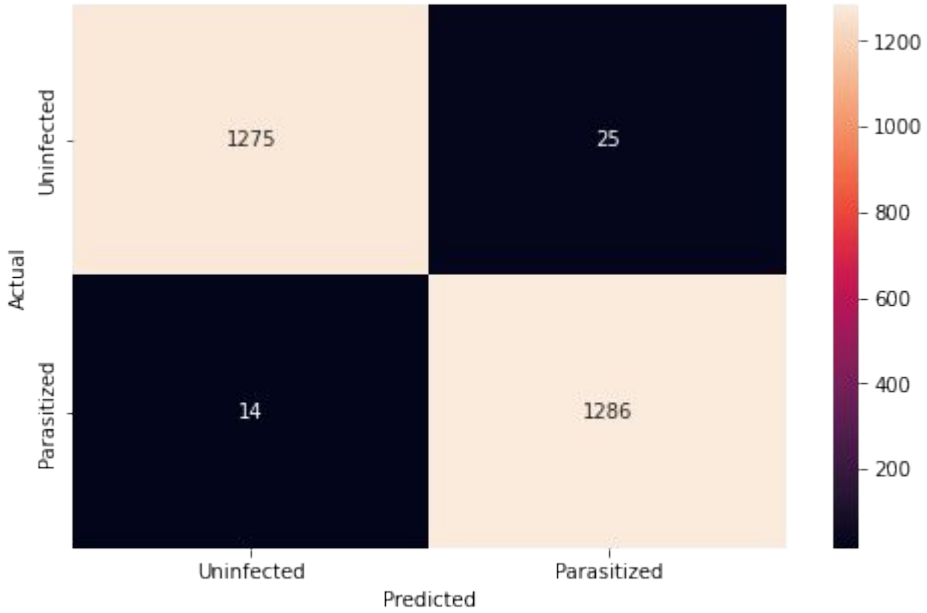
Solution Goals

- The goal of creating a machine learning model is to create a system that can receive patient data and make quick, accurate diagnostic in real time.
- Quick diagnostics can limit suffering and save thousands of lives.
- A single model can replace the work of thousands of professionals. This would allow professionals and resources to be allocated to treatment.
- Researchers can spend valuable resources searching for a cure rather than making everyday human diagnostics.

Results Example

Results of a Convolutional Neural Network

	precision	recall	f1-score	support
0	0.99	0.98	0.98	1300
1	0.98	0.99	0.99	1300
accuracy			0.98	2600
macro avg	0.99	0.98	0.98	2600
weighted avg	0.99	0.98	0.98	2600



Recommendations

Key Recommendations

- Assemble a team of well respected data scientists to devise a model that can accurately and consistently predict when a patient's blood cells are infected or healthy.
- Build in house database and computational infrastructure so the model can be developed and improved rapidly without interference from an outside competitor.
- Assemble a corporate-government liaison team that has experience in dealing with healthcare regulatory agencies. Build institutional capacity for private sector - regulator cooperation, focusing on expediting medically significant product release dates.

Recommendations

Benefits

- The company can develop and license a proprietary model that can be used globally.
- Licensing of the model can generate surplus profits and improve company margins.
- The company can improve public image by providing a life saving technology to people in need.
- New investors may be encouraged to invest in publicly traded shares after learning of this new product.

Costs

- Increased staffing costs to hire and maintain a fulltime team of data scientists and regulatory staff in order to maintain and gain approval for the model.
- Build and maintain a local database and computational system that can perform the modeling.
- 150/250k annual cost per additional staff required to fulfill this goal.
- 100k annual cost for building and maintaining the database and computational system.

Recommendations

Key Risks

- The inability to produce a model that can accurately identify when a patient is infected or not.
- Failure in getting the model approved for use in real life diagnostics.
- The model misdiagnosis patients and harm is caused to people.
- A competitor may produce a more accurate model that is chosen to be used.

Further Analysis

- More information needs to be gathered to properly understand if a model can perform better than professional experts.
- More analysis is needed to know the true costs of bringing this product to market.
- Differences in global data may present a potential problem that could be difficult to solve.