

Language Recognition With Machine Learning

Muth, Isaac (School: Mandan High School)

The purpose of my project is to show the feasibility of useful machine learning applications on a small scale and with low performance requirements, allowing it to run on common devices and therefore be accessible to more people. For my demonstration, I wrote, trained, and tested a machine learning program capable of learning to recognize text in various languages by learning from samples of those languages. I chose this path because of the large amount of data available for training and because all training could be done in a controlled environment. After the program was developed, its ability to learn was then tested. I ran two series of tests, one with six languages and one with twelve languages. In both cases, its ability to identify sentence-sized fragments of text was tested. In both cases, it performed well, leveling off at approximately 97.5% accuracy with six languages, and approximately 86.9% accuracy with twelve languages. Furthermore, the trained program remained fairly small in terms of storage required, and computational times also remained low. While more testing is needed, it seems likely this lightweight approach could also be applied to other problems. The combination of low performance requirements, rapid training time, and accuracy show that machine learning is practical on common devices.