

Data Visualization: Final project

RNN Occlusion visualization

Reid McIlroy-Young

June, 2018

Introduction

The existing work on visualization of recurrent neural networks (RNN) is quite lacking, the best I am aware of is a sequence to sequence and word/character model visualizer that uses a parallel coordinates plot: lstm.seas.harvard.edu/client/index.html. This is a visual of how the Long Short Term Memory (LSTM) layers are working but is busy, complex and unintuitive. There have been a few other pieces of work done on the topic, mostly focusing on sequence to sequence or generative models, but none on classifiers that I could find. The goal of this project is get a better visualizer of LSTM/RNN classifiers

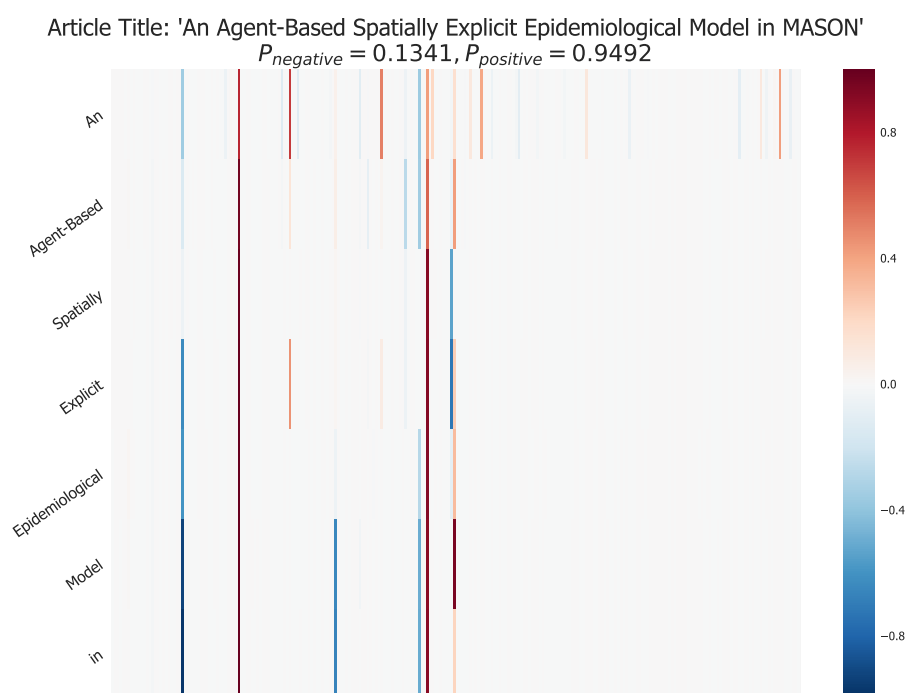


Figure 1: Example of output activations for a positive example