

Proposal: Create a machine learning model to identify authentic vs fake comments for FCC Net Neutrality.

Problem:

The FCC Net Neutrality public comment section has been spammed by bots with fake pro-repeal comments.

Client:

The internet using public and the FCC.

Data:

A SQL data base of public comments with their text and the associated zip codes.

Approach:

Develop a model on a sample of the 161GB of data expanding upon on the work of [Jeff Keo's post at hackermoon.com](#), to create a simple classifier to identify which texts are authentic.

Deliverables:

A report with a machine learning model that can distinguish between