Wildfire Cause Prediction with GNNs

Jimmy Howerton & Mark Tenzer

OUTLINE

01
MOTIVATION
Why study wildfires?

02
DATA
Fires, Populations, Weather

03
METHODS
Baselines, GCN, GraphSAGE, GIN

04
PROGRESS
Data subset results

05
PLAN
Full dataset results, comparisons, reasoning

Why Predict Wildfire Cause?

- Expense of wildfire cause evaluation
- Graphical Approach
- Data augmentation

KNOWN FACTORS

POPULATION

Arson and other human-caused fires usually occur near population centers.

ELEVATION

Lightning is much more likely at high elevations.

WEATHER

Naturally-caused fires, on average, require drier conditions.

Also, thunderstorms predict lightning fires.

DATA

	ROWS	FEATURES
WILDFIRES	1.88M	29

KAGGLE WILDFIRE DB

- All U.S. wildfires, 1992-2015
- Size, Discovery Date, Containment Date, Location, etc.
- Cause

WILDFIRE DATASET

DATA

	ROWS	FEATURES
WILDFIRES	1.88M	29
POPULATION	141K	7

POPULATION DATASET

WORLD CITIES DATABASE

- Countries, Cities, Regions
- Location
- Population

DATA

	ROWS	FEATURES
WILDFIRES	1.88M	29
POPULATION	141K	7
WEATHER	1.56M	32

NOAA GSOD

- 1929-Present
- One reading per day per station
- Missed readings
- Snow, Hail, Wind, etc.
- Temperature, Thunder, Precipitation, Dew Point

BASELINE - MLP

MLP

- 3 Hidden dense layers
  - 64 Neurons per layer
  - ReLU activation
  - Batch normalization
- dense 11-node softmax output

BASELINE - MLP

%53.8

Original Features

BASELINE - MLP

%53.8	Original Features
%50.6	Augmented Features

BASELINE - MLP

%53.8

Original Features

%50.6

Augmented Features

%57.0

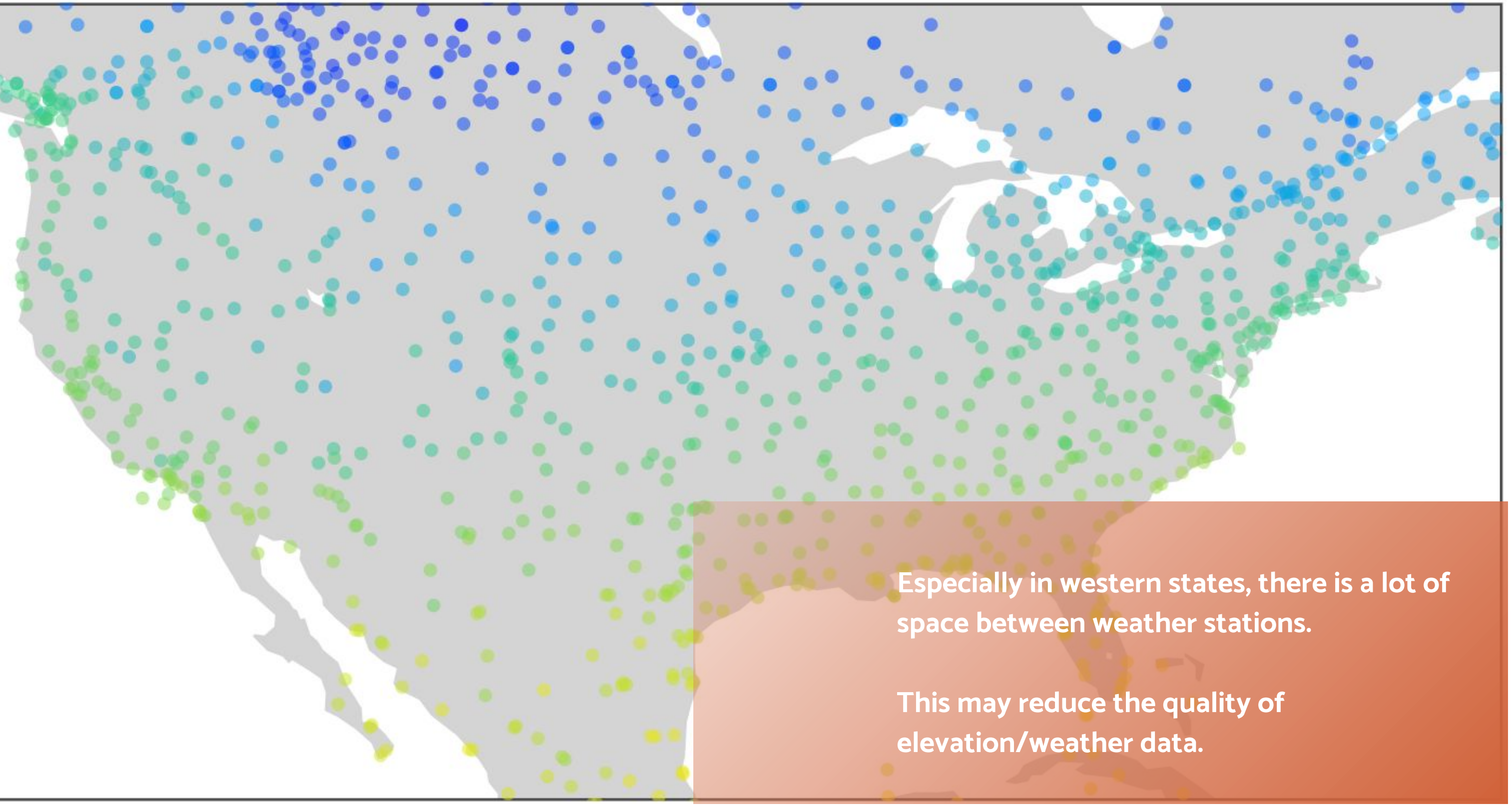
All Features

BASELINE - MLP

%53.8	Original Features
%50.6	Augmented Features
%57.0	All Features

Kaggle
%58.0
(Random forest)

WHY?



Especially in western states, there is a lot of space between weather stations.

This may reduce the quality of elevation/weather data.

Map courtesy of “ircus” at Kaggle.com

CONNECTIONS

SPACE

Connect fires by latitude & longitude

CONNECTIONS

METHODS

SPACE

Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

CONNECTIONS

SPACE

Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

SPACE + TIME

Connect fires by both latitude, longitude, and day of year

CONNECTIONS

SPACE

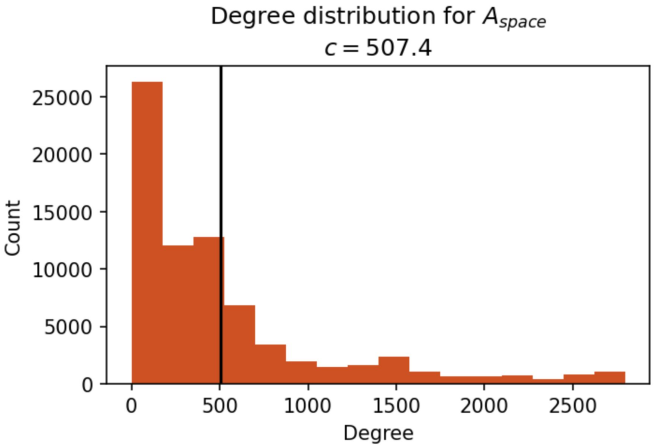
Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

SPACE + TIME

Connect fires by both latitude, longitude, and day of year



CONNECTIONS

METHODS

SPACE

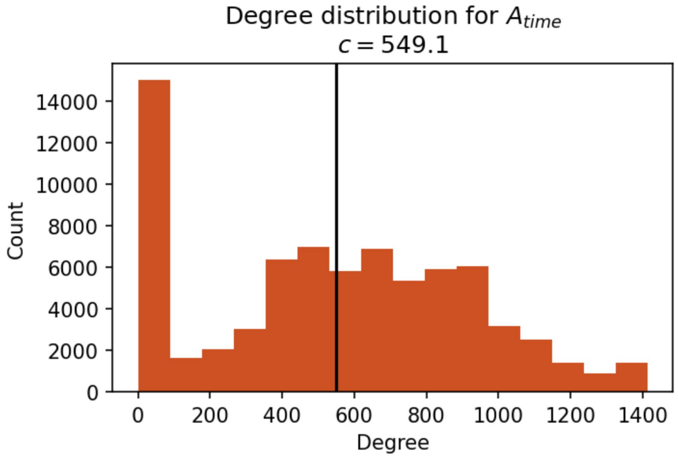
Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

SPACE + TIME

Connect fires by both latitude, longitude, and day of year



CONNECTIONS

METHODS

SPACE

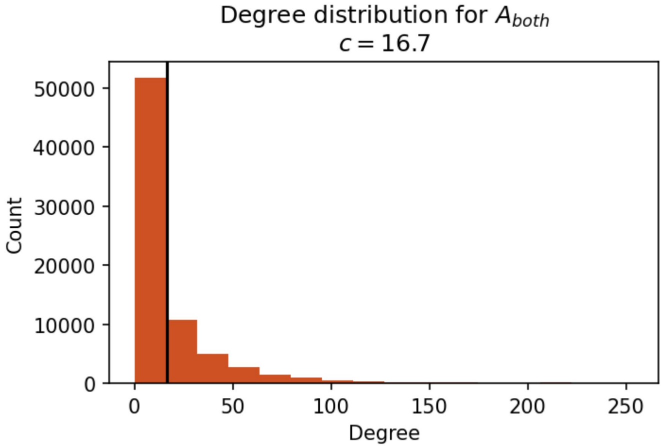
Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

SPACE + TIME

Connect fires by both latitude, longitude, and day of year



METHODS

CONNECTIONS

SPACE

Connect fires by latitude & longitude

TIME

Connect fires by day of year occurred

SPACE + TIME

Connect fires by both latitude, longitude, and day of year

MODELS

GCN

Mean Pooling

METHODS

CONNECTIONS

SPACE
Connect fires by latitude & longitude

TIME
Connect fires by day of year occurred

SPACE + TIME
Connect fires by both latitude, longitude, and day of year

MODELS

GCN
Mean Pooling

GraphSAGE
Max Pooling

CONNECTIONS

SPACE
Connect fires by latitude & longitude

TIME
Connect fires by day of year occurred

SPACE + TIME
Connect fires by both latitude, longitude, and day of year

MODELS

GCN
Mean Pooling

GraphSAGE
Max Pooling

GIN
Sum Pooling

MLP: 1 hidden layer, 64 neurons

CONNECTIONS

METHODS

SPACE
Connect fires by latitude & longitude

TIME
Connect fires by day of year occurred

SPACE + TIME
Connect fires by both latitude, longitude, and day of year

MODELS - DNN Variant

GCN
Mean Pooling

GraphSAGE
Max Pooling

GIN
Sum Pooling

MLP: 1 hidden layer, 64 neurons

Spektral

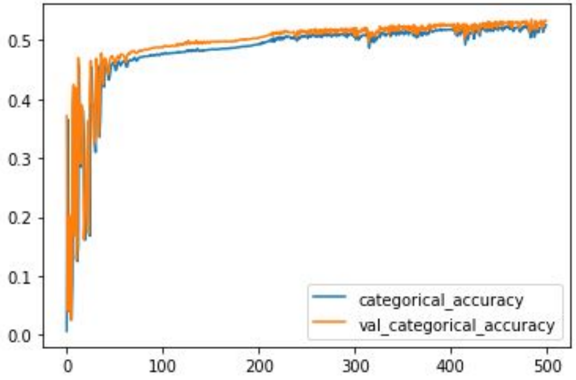
- Python library for graph deep learning
- based on Keras and Tensorflow
- all models run on GPU in Ubuntu
- Implements GCN, GraphSAGE, and GIN into Keras model layers

MODEL SPECS

5	**GNN Layers**
500	**Epochs**
3	**DNN Layers**
Mean	**Imputation**

SPACE + TIME (2015)

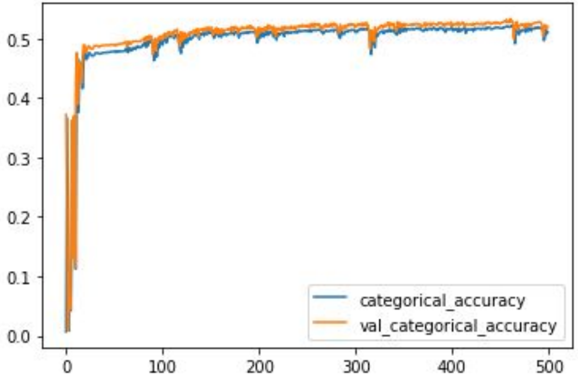
PROGRESS - GCN



GCN

~53.4%

Converges quickly, but performs worse than a simple fully-connected model

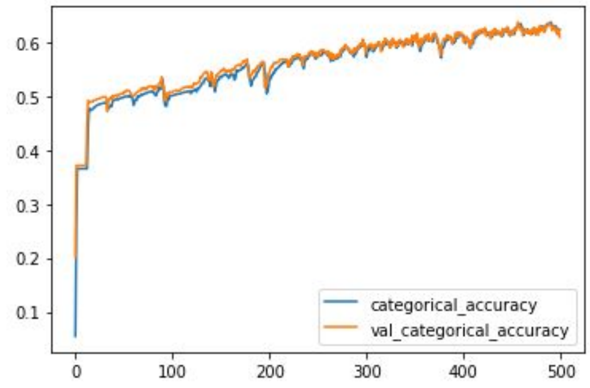


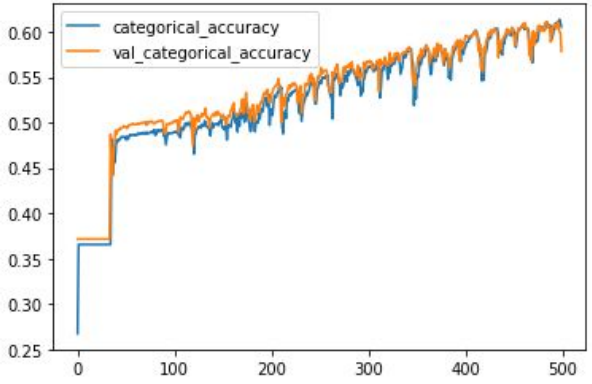
GCN-DNN

~52.0%

Adding the fully-connected model, with the node embedding as input, does not improve performance

SPACE + TIME (2015)





GraphSAGE

~63.0%

Achieves state-of-the-art performance!

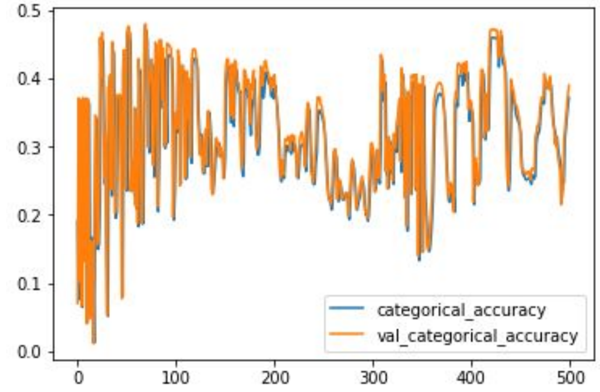
GraphSAGE-DNN

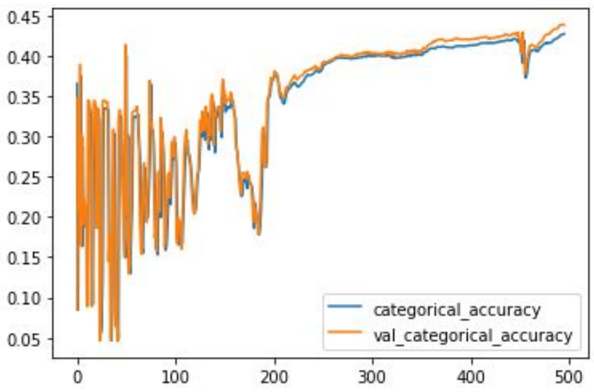
~61.0%

Also outperforms the baseline, but the added DNN does not improve the model.

PROGRESS - GraphSAGE

SPACE + TIME (2015)





GIN

~ ?

GIN struggles to converge after even 500 epochs of training.

GIN-DNN

~43.0%

The addition of a DNN stabilizes the training, but overall accuracy still is poor.

PROGRESS - GIN

PLAN

Proposal

Determining problem, datasets, methods

PLAN

Progress Report

Data augmentation,
Familiarizing with tools

Proposal

Determining problem,
datasets, methods

PLAN

Proposal

Determining problem, datasets, methods

Progress Report

Data augmentation, Familiarizing with tools

Today

Presentation

Running baseline & subset models

PLAN

Proposal
Determining problem,
datasets, methods

Progress Report
Data augmentation,
Familiarizing with tools

Today

Presentation
Running baseline &
subset models

Paper Completion
Run models on full dataset,
compare results,
suggest future work

PLAN

Proposal
Determining problem,
datasets, methods

Progress Report
Data augmentation,
Familiarizing with tools

Today

Presentation
Running baseline &
subset models

Paper Completion
Run models on full dataset,
compare results,
suggest future work

Paper Submission
State of the art results
could make conference
submission worthwhile

Future work

- Run analyses on all 1.88M fires
- Experiment with new NN structures?
- Experiment with other adj. matrices?
- Reduction to binary classification
    - Arson is the most important case!

CREDITS

- https://commons.wikimedia.org/wiki/File:Deerfire_high_res_edit.jpg
- https://static01.nyt.com/images/2019/06/11/us/00waspnest-01/merlin_156060048_755263d2-7783-4275-ae26-583f7f568253-superJumbo.jpg
- https://www.nationalgeographic.com/content/dam/science/2019/10/09/california-wildfire/calif-wildfire-ap_18250148901527.adapt.1900.1.jpg
- https://spektral.graphneural.network/

- Presentation template by Slidesgo
- Icons by Flaticon
- Infographics by Freepik
- Images created by Freepik

THANK YOU

Does anyone have questions?