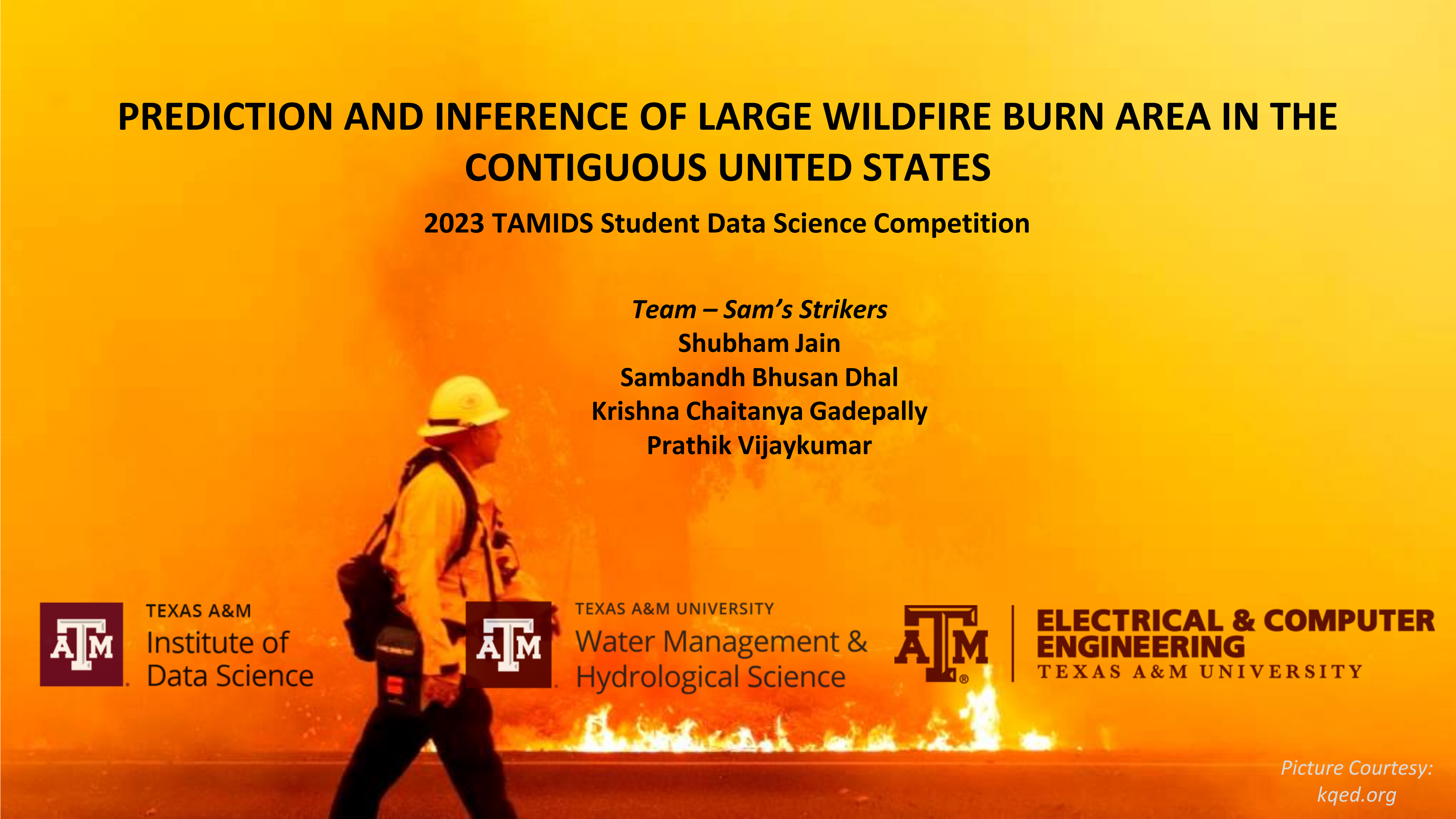


PREDICTION AND INFERENCE OF LARGE WILDFIRE BURN AREA IN THE CONTIGUOUS UNITED STATES

2023 TAMIDS Student Data Science Competition

Team – Sam's Strikers

Shubham Jain

Sambandh Bhusan Dhal

Krishna Chaitanya Gadepally

Prathik Vijaykumar

TEXAS A&M Institute of Data Science

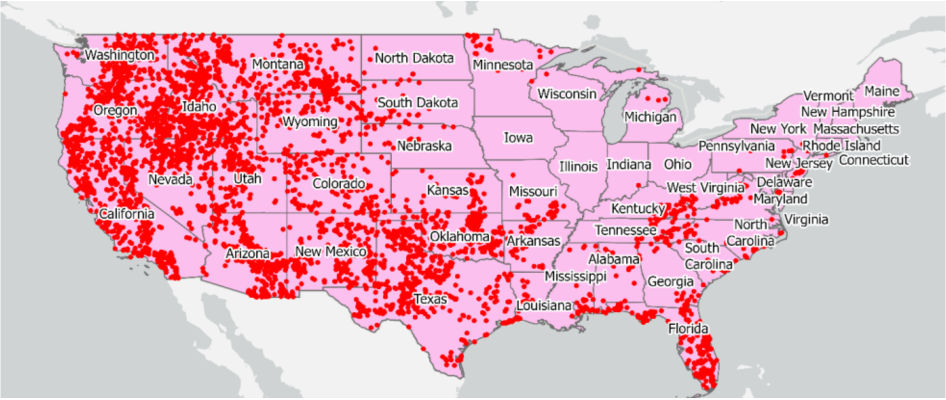
TEXAS A&M UNIVERSITY Water Management & Hydrological Science

ELECTRICAL & COMPUTER ENGINEERING TEXAS A&M UNIVERSITY

Picture Courtesy: kqed.org

Introduction

- Primary focus on two challenges: **wildfire behavior** prediction and effective communication of research findings
- Wildfire behavior is complex, and predicting it accurately is challenging
- Effective communication of research findings to end-users is critical in addressing the wildfire problem
- In last decade, the US witnessed **70,000** wildfires per year, burning **7 million acres** of land on average
- Texas had **3,700** wildfires in 2021, burning around **200,000** acres of land
- Data-driven approaches can help in predicting wildfire spread and ultimately reducing their impact
- A **Deep Learning** approach is used to predict burn area for large wildfire occurrences based on **climate forcings** and **geological characteristics**

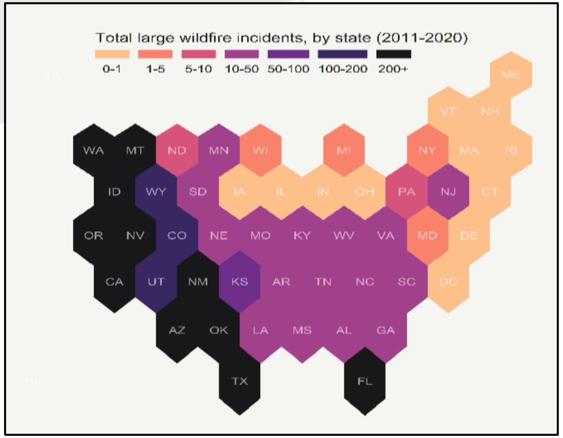


Map of large wildfire incidents in the contiguous US between 2011 and 2020 (4538 incidents)

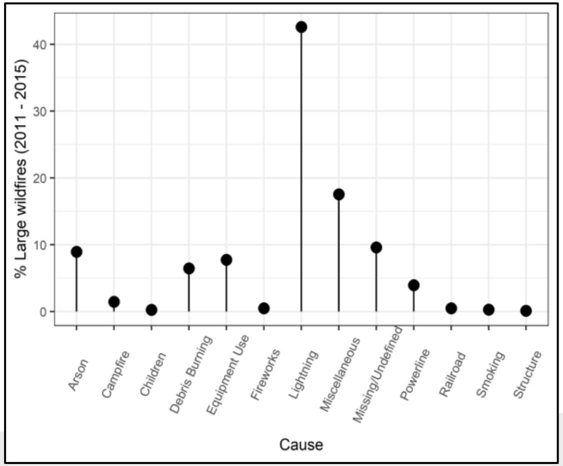
Wildfire Data

- GIS data for large wildfire locations and burn area boundaries downloaded from **Monitoring Trends in Burn Severity Program**.
- Large wildfire thresholds – **1000** acres in Western US and **500** acres in Eastern US.
- 10 years of data collected **(2011 – 2020)**
- **4,538** incidents covering **87,305 square miles** of burn area used in analysis

Additional wildfire data obtained from **Fire Program Analysis Fire-Occurrence Database (FPA-FOD)** for information on wildfire causes



Large wildfire incidents by state

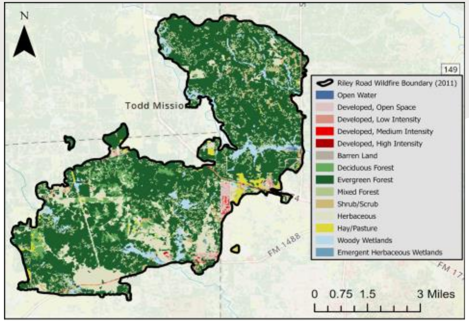


Major causes of large wildfires in US

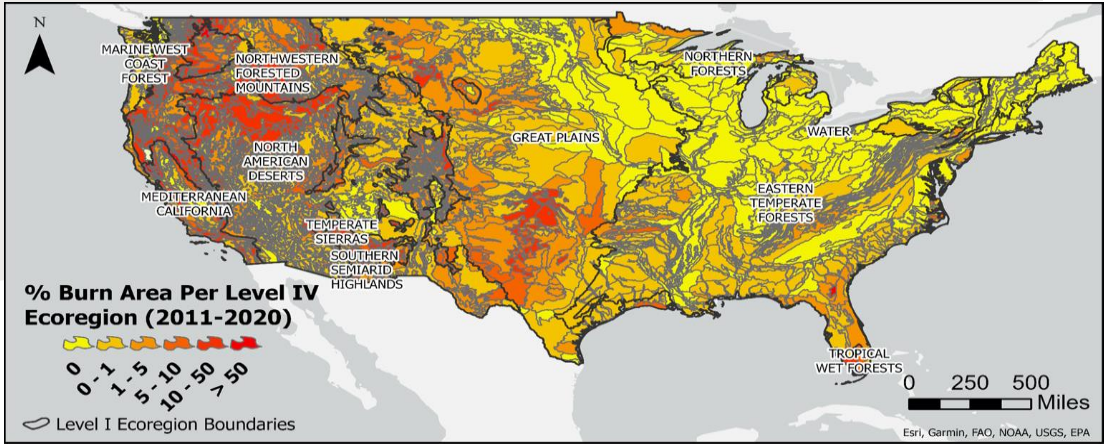
Meteorological and Topography Data

	Climate	**PRISM**	Precipitation, Temperature, Vapor Pressure Deficit (Minimum and Maximum)
		GRIDMET	Palmer Drought Severity Index (PDSI), Potential Evapotranspiration (PET)
	Land cover	**National Land Cover Database (NLCD 2016)**	Open Water, Developed, Barren, Forests, Shrub/Scrub, Hay/Pasture, Cultivated Crops, Wetlands
	MODIS	**MOD13A3 Version 6**	Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI)
	Topography	**USGS DEM**	Elevation (meters)
	Ecoregion	**US EPA Ecoregions**	Level I and Level IV ecoregion boundaries

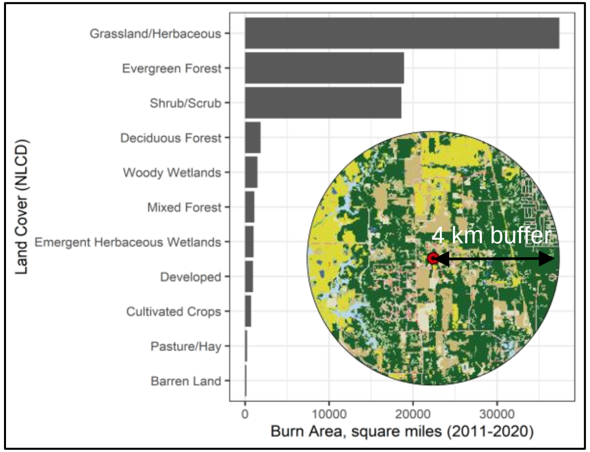
Data Exploration



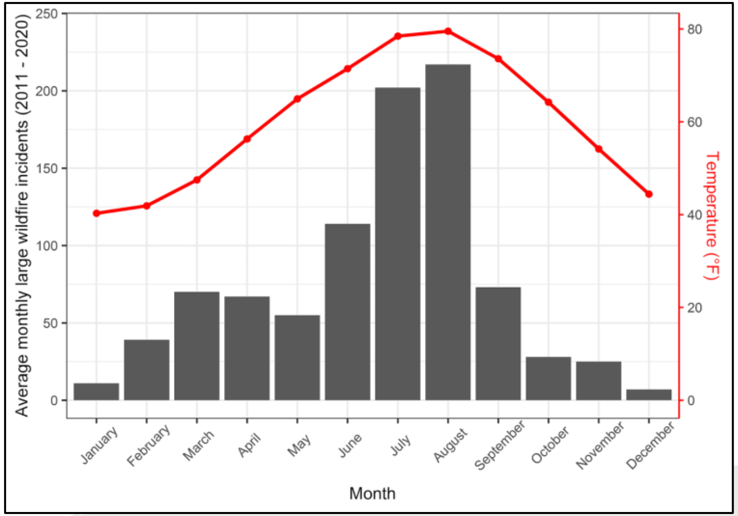
NLCD land cover for area burned in Riley Road wildfire northwest of Houston burning 19,000 acres of land



Percent of Level IV Ecoregion land burned in large wildfires between 2011 - 2020



Burn area by NLCD land cover in large wildfires between 2011-2020

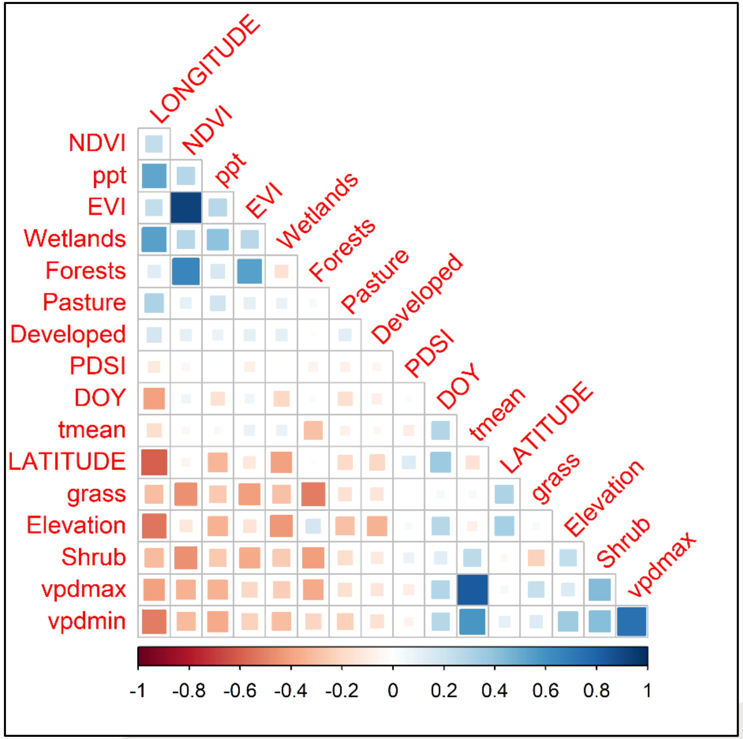


Average monthly large wildfires (bar) in the contiguous US and the mean monthly temperature (line).

Predicting Burn Area

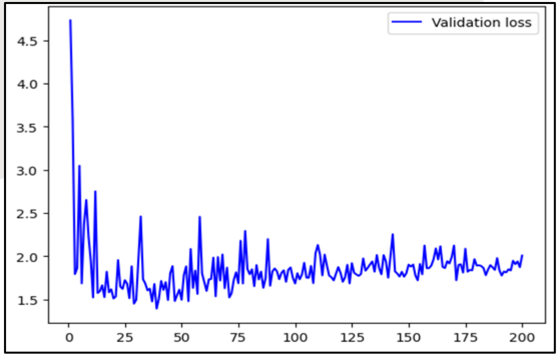
Features gathered for 4,536 large wildfires

Feature	Description	Min	Max
LATITUDE	Latitude coordinates of wildfire occurrence (decimal degrees)	25.2	49
LONGITUDE	Longitude coordinates of wildfire occurrence (decimal degrees)	-124.1	-72.8
DOY	Wildfire ignition day of year	1	365
ppt	Total monthly precipitation for month of wildfire ignition	0	1063.2
tmean	Average monthly temperature for month of wildfire ignition	-5.3	36.8
vpdmax	Maximum vapor pressure deficit for month of wildfire ignition	2.7	81.8
vpdmin	Minimum vapor pressure deficit for month of wildfire ignition	0	35.3
PDSI	Palmer Drought Severtiy Index during ignition date	-8.1	7.6
Developed	% NLCD developed around 4-kilometer buffer of wildfire ignition	0	64.2
Forests	% NLCD forests around 4-kilometer buffer of wildfire ignition	0	99.8
Shrub	% NLCD shrub/scrub around 4-kilometer buffer of wildfire ignition	0	100
grass	% NLCD grasslands/herbaceous around 4-kilometer buffer of wildfire ignition	0	100
Pasture	% NLCD hay/pasture around 4-kilometer buffer of wildfire ignition	0	74
Wetlands	% NLCD wetlands around 4-kilometer buffer of wildfire ignition	0	100
NDVI	Normalized Difference Vegetation Index for month of wildfire occurrence	0.1	0.9
EVI	Enhance Vegetation Index for month of wildfire occurrence	0	0.7
Elevation	Elevation of wildfire occurrence	-2	3507

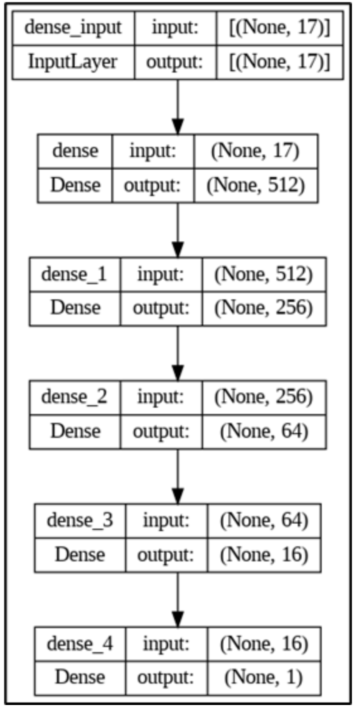


Correlation plot for features used in the study

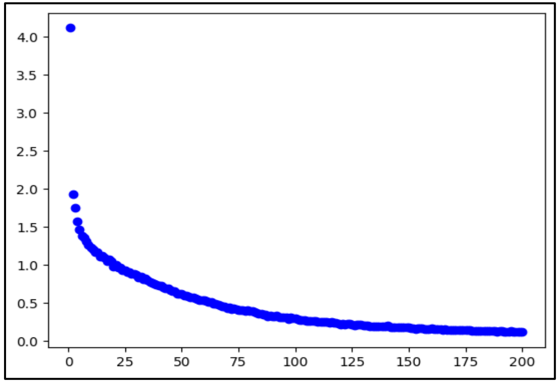
Modeling and Analysis



Validation loss for the DNN model



DNN architecture used in the study



Training loss for the DNN model

$$Error\ rate = \frac{\sum_1^N \left| y_{obs} - y_{pred} \right|}{\sum_1^N \left| y_{obs} \right|}$$

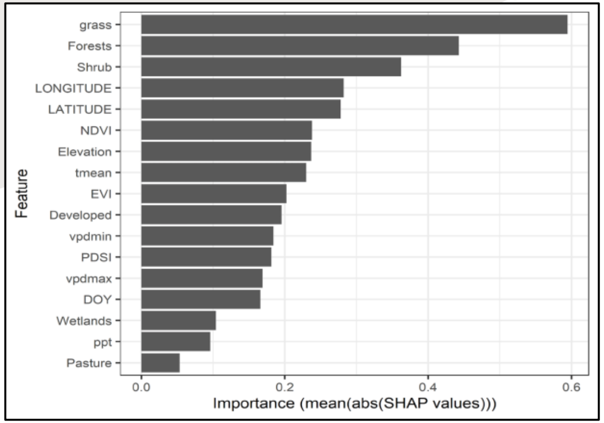
Model Specifications:

- Train/Test split ratio – **80/20**
- Number of repeats: **3**
- Activation function: **ReLU**
- Learning rate: **0.001**
- Batch size: **32**
- Number of Epochs: **200**
- Test Error (MSE): **0.055** to **0.06**

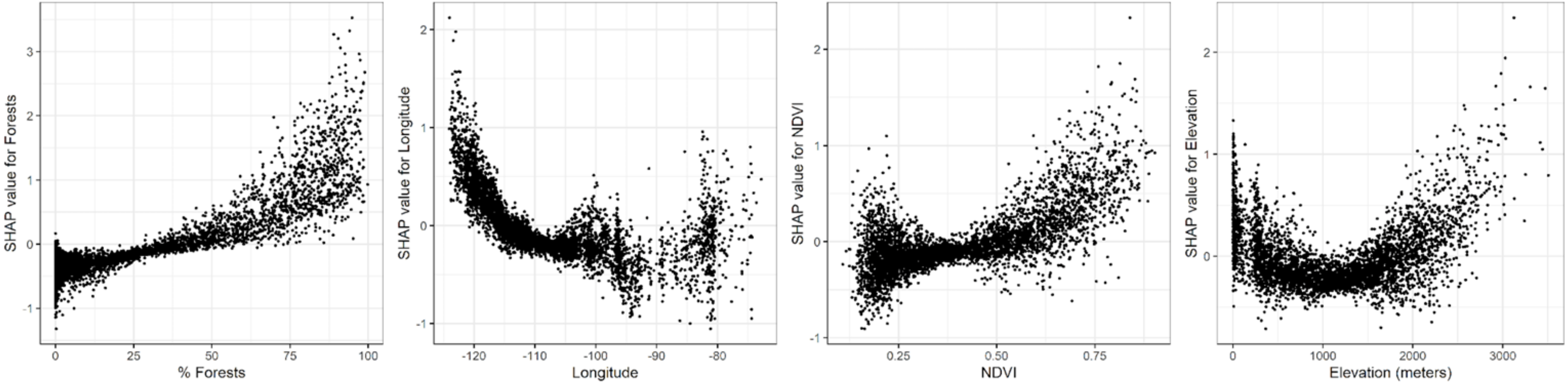
Model Inference

- Using **shapley values** (from game theory) for feature interactions and importance
- Positive SHAP value indicated increase in model's prediction due to the feature in analysis, and vice-versa for negative values
- For any observation -

  Model prediction = Average prediction + sum of all SHAP values



Feature importance in the DNN model obtained from SHAP values



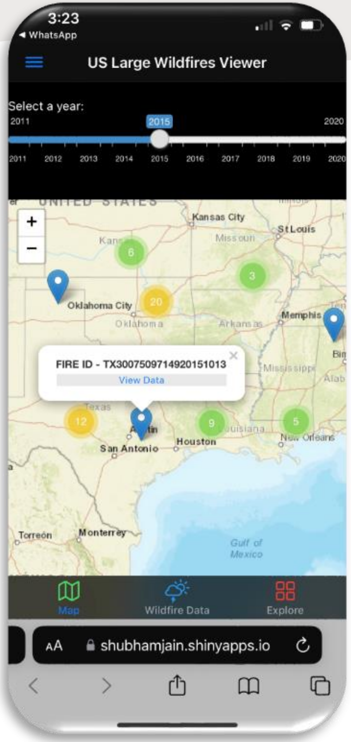
Partial Dependence Plots showing the interactions between features and burn area using SHAP values

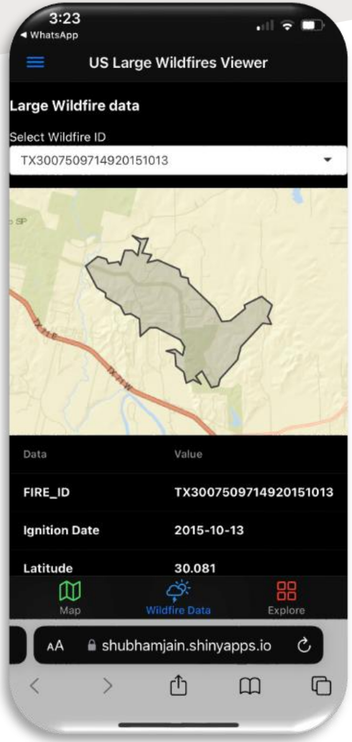
Conclusion and Recommendations

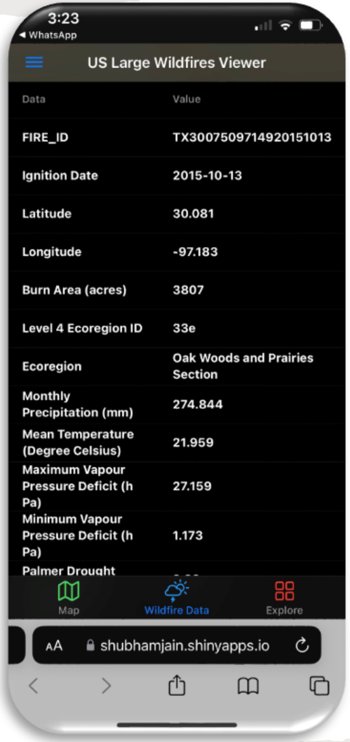
- The **DNN** model performed reasonably well with an average test MSE of **0.055**
- **Land cover** and **location of wildfire occurrence** were most influential in determining burn area
- Future work could involve learning the effects of other features such as **soil characteristics** and **dead fuel moisture** on burn area
- **Data collection, modeling** and **on-ground action** need to be looked together for an effective strategy to mitigate impact of forest fires
- Promoting **interdisciplinary research** with more collaboration between Natural Sciences and Data Science
- A prototype interactive web-tool to visualize wildfire data and characteristics for end-users
https://shubhamjain.shinyapps.io/Wildfires/

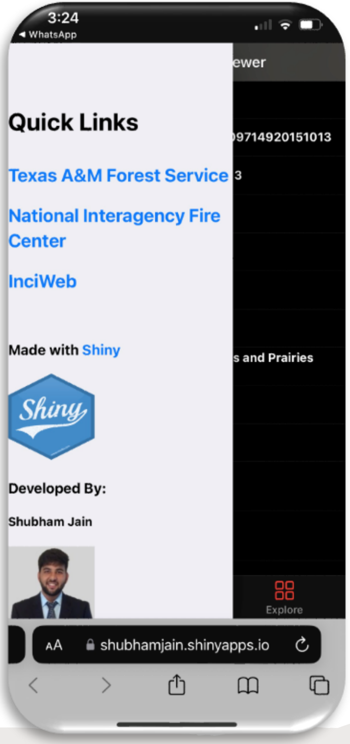
Mobile friendly data visualization tool

Web tool


3:23
WhatsApp
US Large Wildfires Viewer
Select a year:
2015
2011 2012 2013 2014 2015 2016 2017 2018 2019 2020
Kansas City
St Louis
Oklahoma City
Memphis
FIRE ID - TX3007509714920151013
View Data
San Antonio
Houston
New Orleans
Torreon
Monterrey
Gulf of Mexico
Map
Wildfire Data
Explore
shubhamjain.shinyapps.io


3:23
WhatsApp
US Large Wildfires Viewer
Large Wildfire data
Select Wildfire ID
TX3007509714920151013
Data
Value
FIRE_ID
TX3007509714920151013
Ignition Date
2015-10-13
Latitude
30.081
Map
Wildfire Data
Explore
shubhamjain.shinyapps.io


3:23
WhatsApp
US Large Wildfires Viewer
Data
Value
FIRE_ID
TX3007509714920151013
Ignition Date
2015-10-13
Latitude
30.081
Longitude
-97.183
Burn Area (acres)
3807
Level 4 Ecoregion ID
33e
Ecoregion
Oak Woods and Prairies Section
Monthly Precipitation (mm)
274.844
Mean Temperature (Degree Celsius)
21.959
Maximum Vapour Pressure Deficit (h Pa)
27.159
Minimum Vapour Pressure Deficit (h Pa)
1.173
Palmer Drought
Map
Wildfire Data
Explore
shubhamjain.shinyapps.io


3:24
WhatsApp
Quick Links
Texas A&M Forest Service
National Interagency Fire Center
InciWeb
Made with Shiny
Shiny
Developed By:
Shubham Jain
Explore
shubhamjain.shinyapps.io

Team – Sam’s Strikers

Shubham Jain
PhD Student
Water Management and Hydrological Science

Sambandh Bhusan Dhal
PhD Candidate
Electrical and Computer Engineering

Krishna Chaitanya Gadepally
MS Student
Electrical and Computer Engineering

Prathik Vijaykumar
MS Student
Electrical and Computer Engineering

Hearty gratitude:

Chevron

TEXAS A&M UNIVERSITY
Statistics

TEXAS A&M
Institute of
Data Science

ELECTRICAL & COMPUTER
ENGINEERING
TEXAS A&M UNIVERSITY

TEXAS A&M UNIVERSITY
School of Performance,
Visualization & Fine Arts

Thank You!

Questions?