University of Nebraska-Lincoln
Doctoral Dissertation Defense

Applications of Artificial Intelligence on Drought Impact Monitoring and Assessment

Beichen Zhang

May 7th, 2024

Lincoln, NE

2017 Northwest A&F University (China)
Bachelor of Science
Geographic Information Science

2019 University of Nebraska-Lincoln
Master of Science
Natural Resource Sciences
(Climate Assessment and Impacts)

NATIONAL DROUGHT MITIGATION CENTER
UNIVERSITY OF NEBRASKA

2024 University of Nebraska-Lincoln
PhD Candidate
Natural Resource Sciences
(Climate Assessment and Impacts)
Advisors: Dr. Tsegaye Tadesse
Dr. Michael J. Hayes

THE DAUGHERTY WATER for FOOD GLOBAL INSTITUTE at the University of Nebraska

Dissertation Committee

Advisors

Chair
Tsegaye Tadesse, Ph.D.
Research Professor
Applied Climatologist
Remote Sensing Expert
School of Natural Resources

Co-Chair
Michael Hayes, Ph.D.
Professor
Climatologist
School of Natural Resources

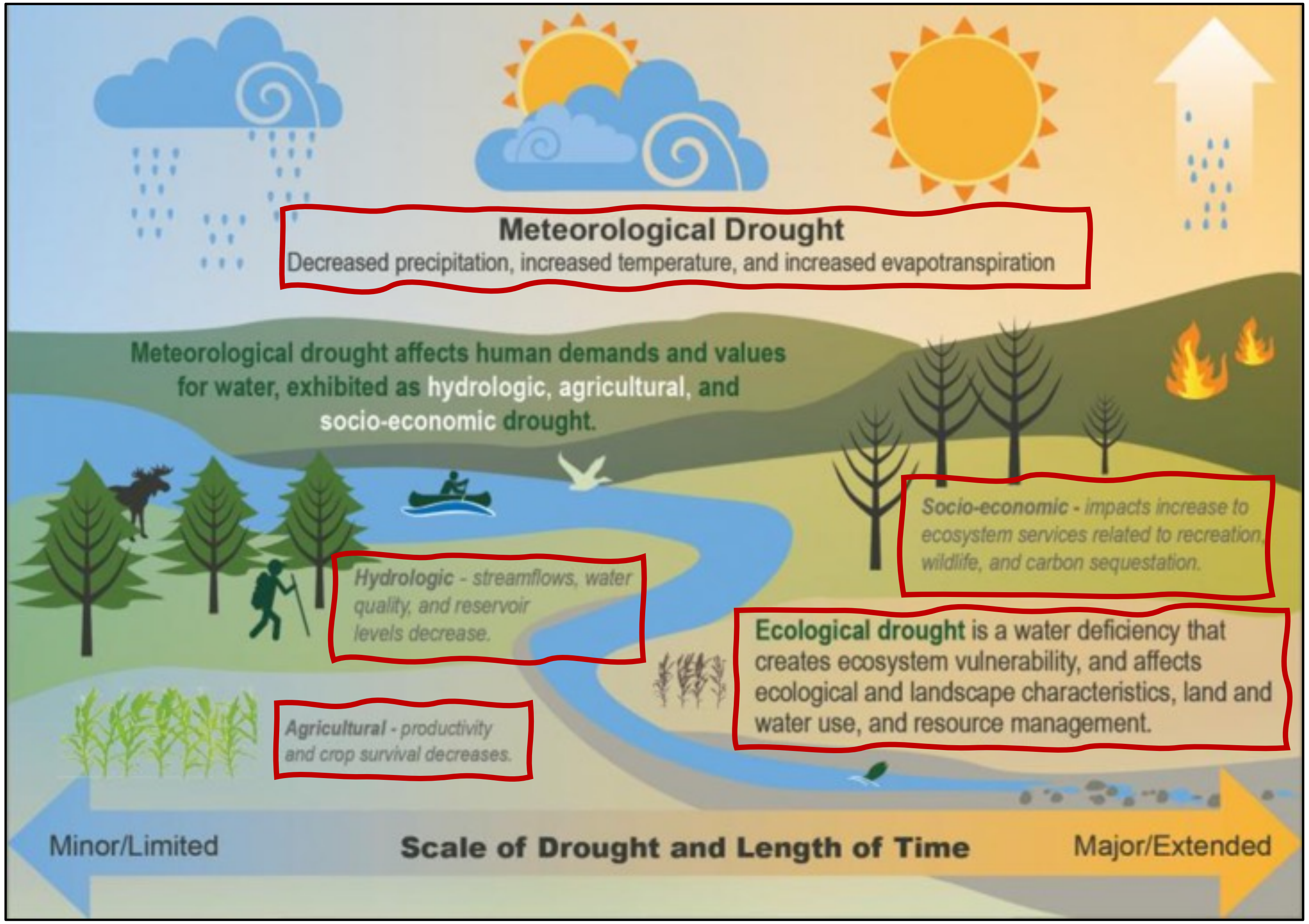
Members

Brian Wardlow, Ph.D.
Professor
Remote Sensing Scientist
School of Natural Resources

Xueheng Shi, Ph.D.
Assistant Professor
Statistician
Department of Statistics

Outside Representative
Ashok Samal, Ph.D.
Professor
Computer Scientist
School of Computing

What is Drought?



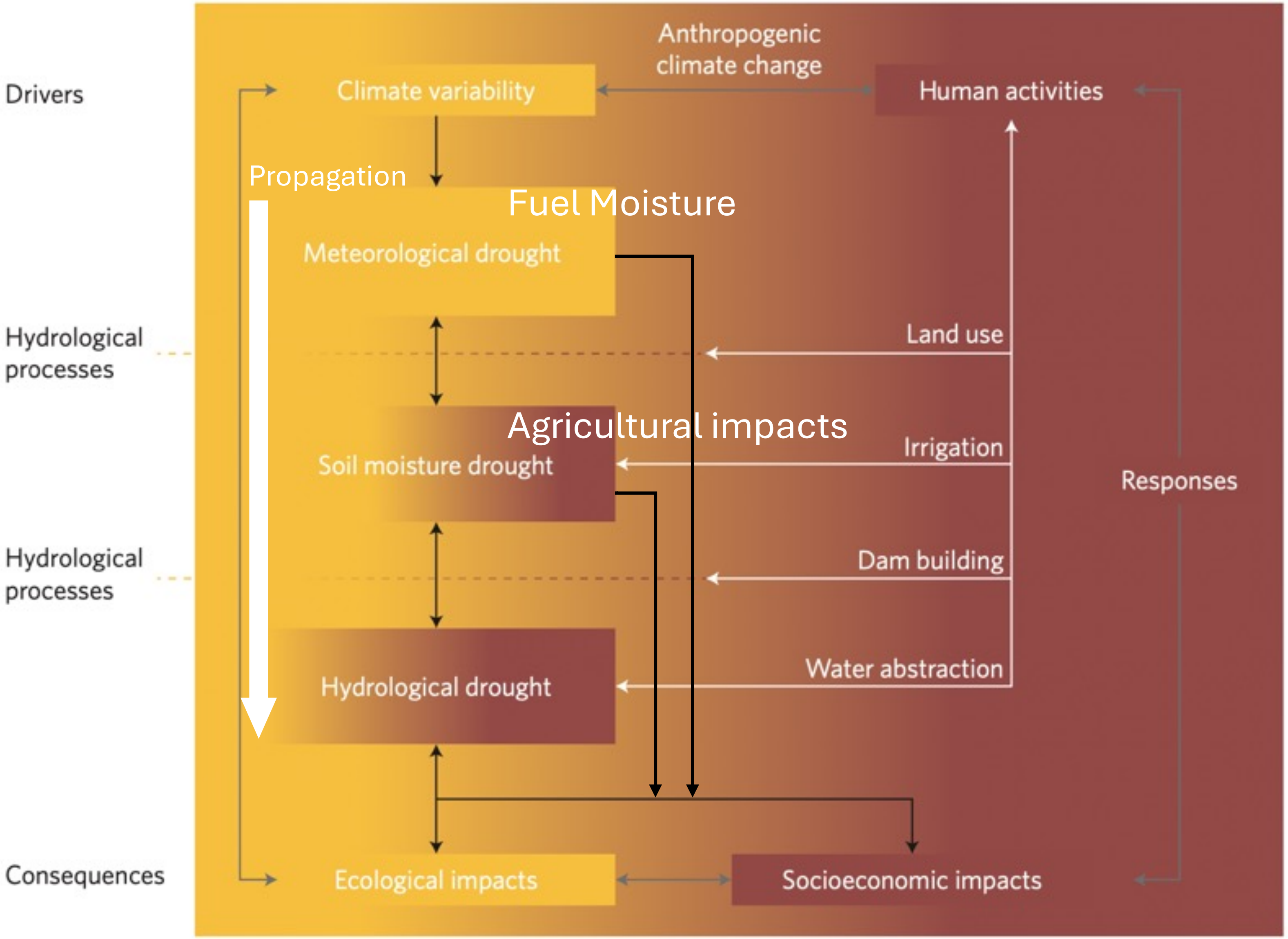
(USDA, 2017)

Wilhite and Glantz (1985):
"A long period with no rain, especially <u>during a planting season</u>"
"An extended period of dry weather, especially <u>one injurious to crops</u>"

- Meteorological Drought
- **Agricultural Drought**
- Hydrological Drought
- **Socioeconomic Drought**
- **Ecological Drought** (Crausbay et al., 2017)
- Flash Drought (Svoboda et al., 2002)

"A deficit between the demand and supply in a stage in the water cycle at a space and time range"

Drought and Drought Impacts



(Van Loon et al., 2016)

Drought and Drought Impacts

(NCEI, 2024)

Drought Impacts

A loss or change related to drought events at a specific place and time.

(NDMC, 2024)

Characteristics of Drought Monitoring and Assessment

Hydrometeorology and Remote Sensing

Probabilistic Properties of Drought
(Operational Definition of Drought)

Intensity

Duration

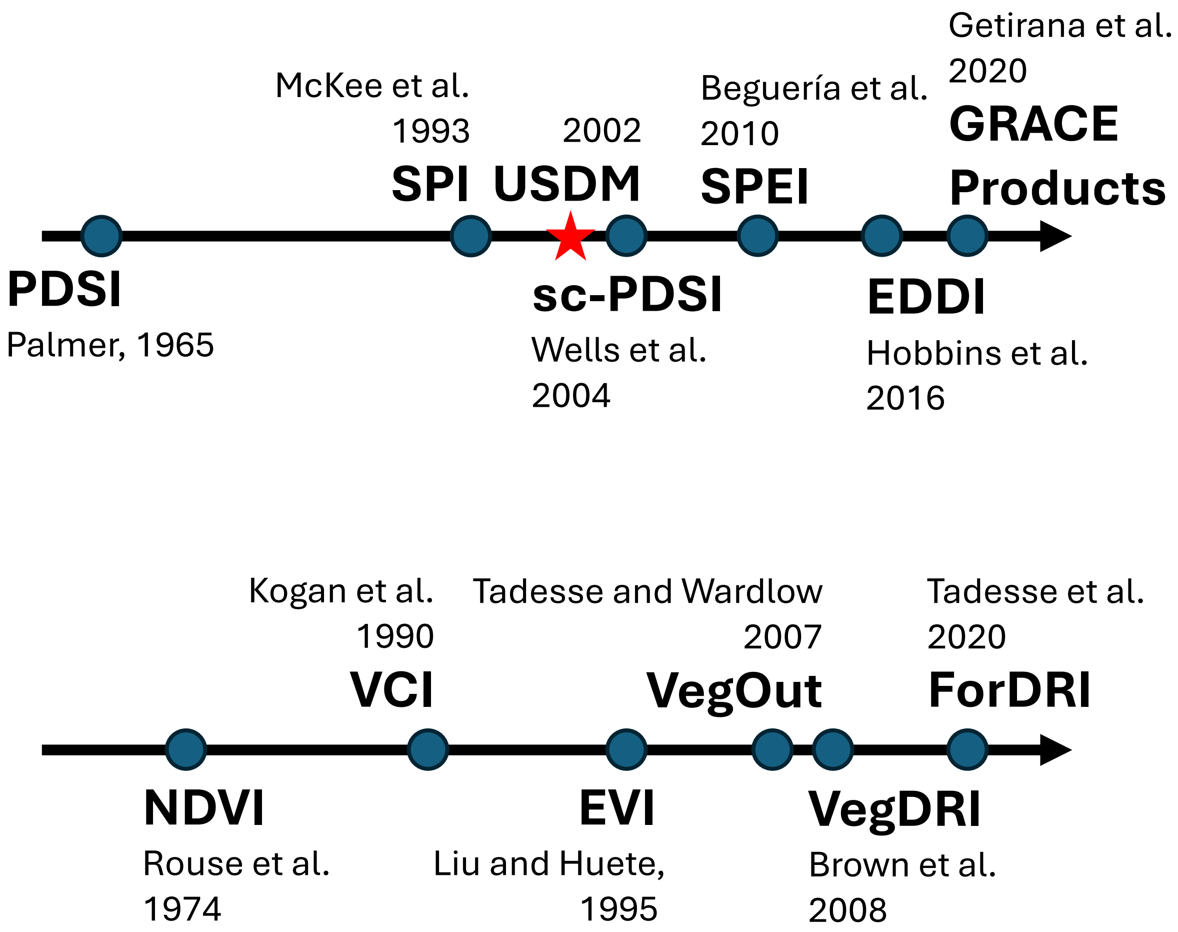
Severity

Frequency

Spatial Extent

(Mishra and Singh, 2011)

Development of Drought Indicators and Indices

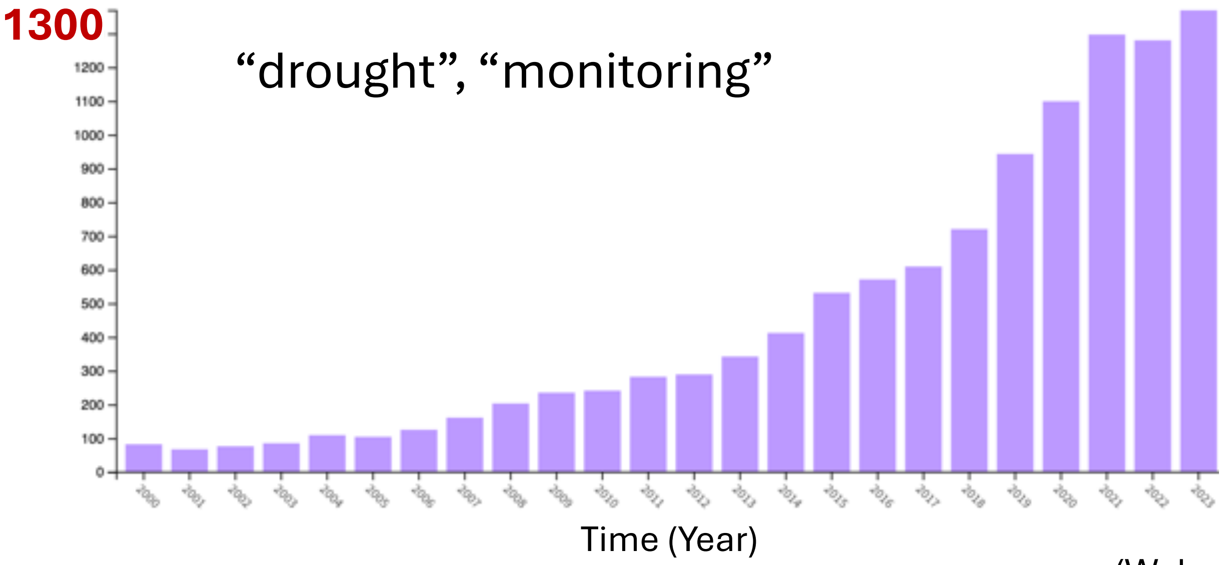


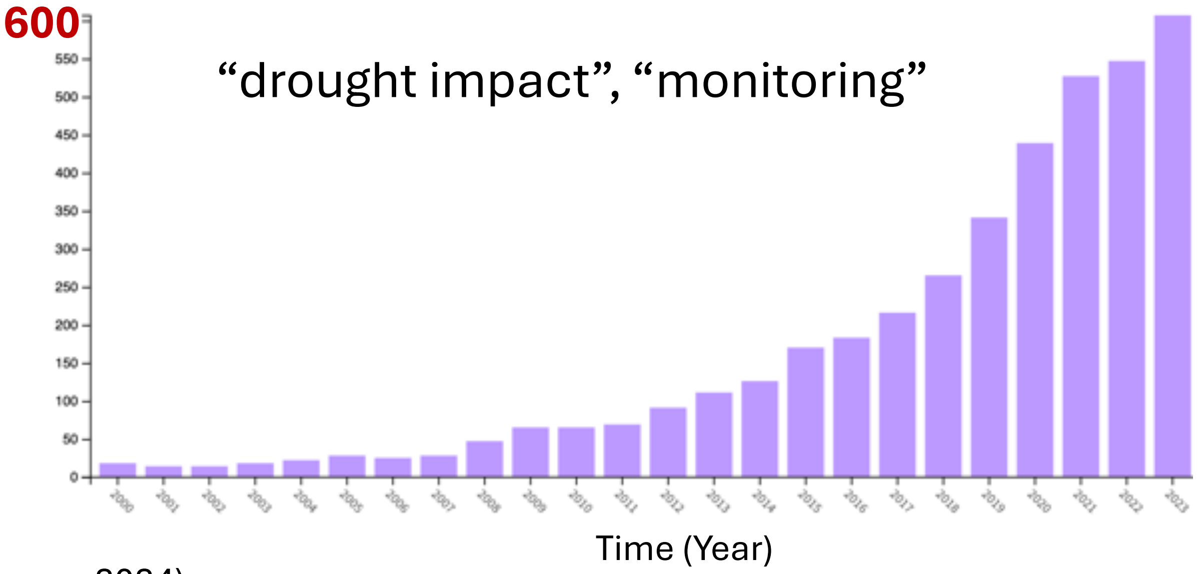
Drought Impact Monitoring and Assessment


Drought indicators revisited: the need for a wider consideration of environment and society

Bachmair et al. 2016:

55 out of 86 studies did not consider drought impact
21 studies focused crop failure and vegetation stress
10 studies (~10%) employed text-based data and other impact variables



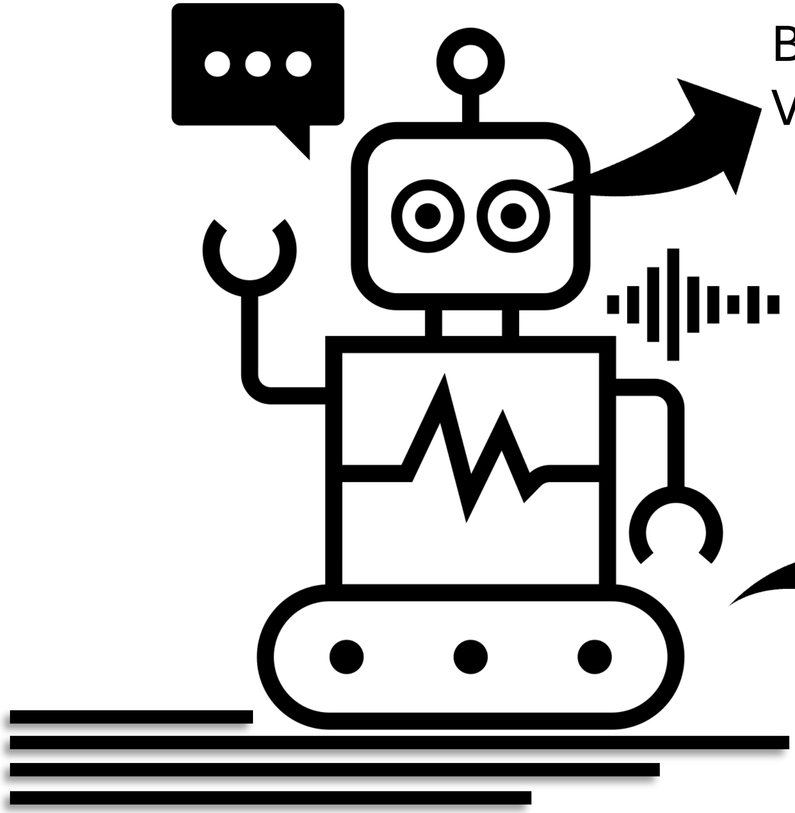


(Web of Science, 2024)

Artificial Intelligence (AI)

Being able to learn from data without hard-coded rules:
Machine Learning (ML), Deep Learning (DL), and more

Being able to talk =>
Natural Language Processing (NLP) and more



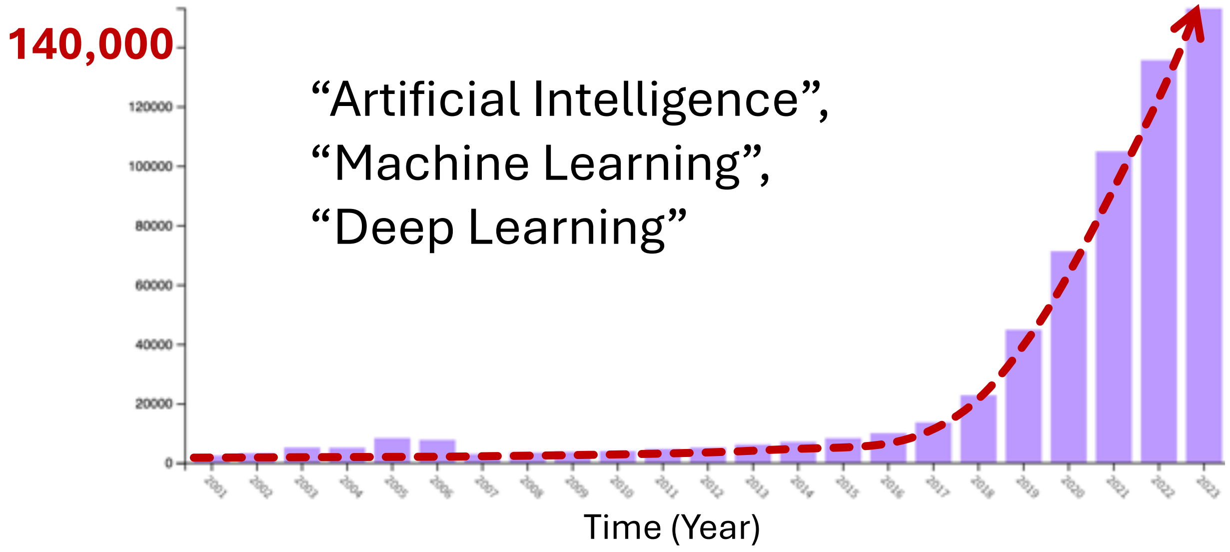
Being able to see => Computer Vision (CV) and more
Vastly used in **remote sensing**

Being able to listen => Speech Recognition and more

Being able to move => Robotics and more

Why Apply AI to Monitor and Assess Drought Impacts?

Rapid development of AI in the past five to ten years.

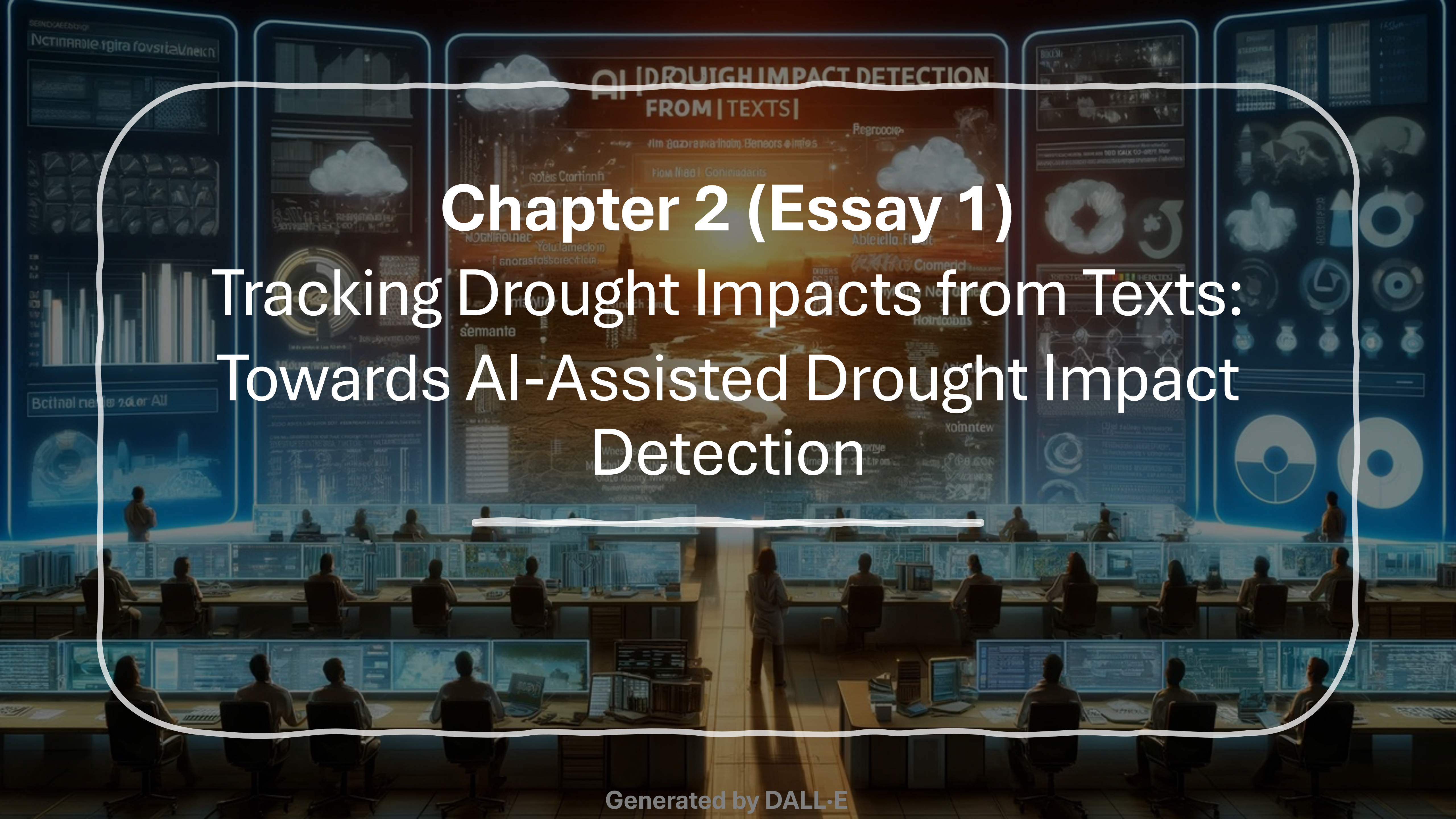


(Web of Science. 2024)

For drought impacts specifically:

- Strong predictive capability
- Non-linear and high-dimensional data
- Flexibility in data types and formats
- Robustness to noise and outliers
- Scalability

	Gaps	AI Methods
Essay 1	Insufficient and constrained quantitative datasets to observe and describe drought impacts beyond agricultural and ecological sectors.	NLP & DL
Essay 2	Understudied links between drought indices and drought impacts on multi-dimensional sectors.	Explainable ML
Essay 3	Particularly, lack of studies on drought impacts on socioeconomic sectors.	Causal ML



Chapter 2 (Essay 1)

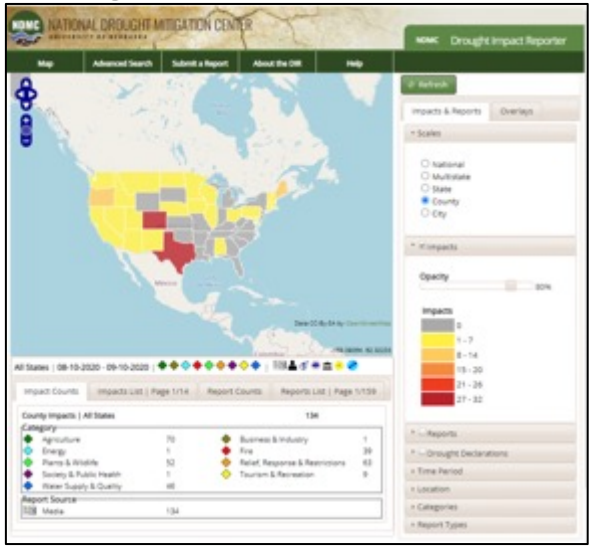
Tracking Drought Impacts from Texts: Towards AI-Assisted Drought Impact Detection

Generated by DALL·E

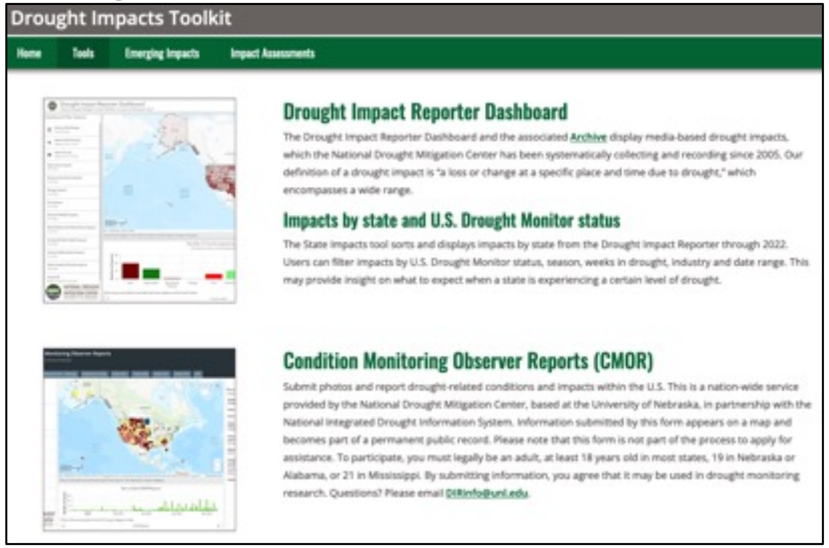
Brief Background and Research Justification

Text-based Observations of Drought Impacts

Drought Impact Reporter (DIR)



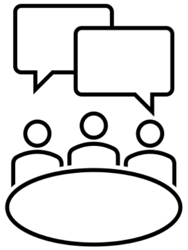
Drought Impacts Toolkit (DIT)



Publications (2020, 2023) from de Brito's research team

- **Textual datasets** have been considered and used to monitor and assess drought impacts.
- DIT provides **high-quality** and **hand-labeled** textual data related to drought awareness and impacts.
- No study yet applies DL-based NLP models to identify drought impacts.

Research Questions



Can we utilize the power of DL-based NLP models to assist users in categorizing drought impacts?

Can we apply the developed NLP models to help enlarge the drought impact database?

Research Objectives

- Fine-tuning and prompt-engineering the DL-based NLP models.
- Transferring the models on social media data.
- Performance comparison with the benchmark model.
- Spatial and temporal analysis of the predicted categories from social media.

NLP models

Radford et al., 2018, 2019;
Brown et al., 2020;
Ouyang et al., 2022

Generative Pre-trained Transformer (GPT)

Decoder

GPT, GPT-2, GPT-3, InstructGPT

GPT-4

Devlin et al., 2018

Bidirectional Encoder Representations from Transformers (BERT)

Encoder

Bert-like Models (Variants)

Robust BERT
Liu et al., 2019

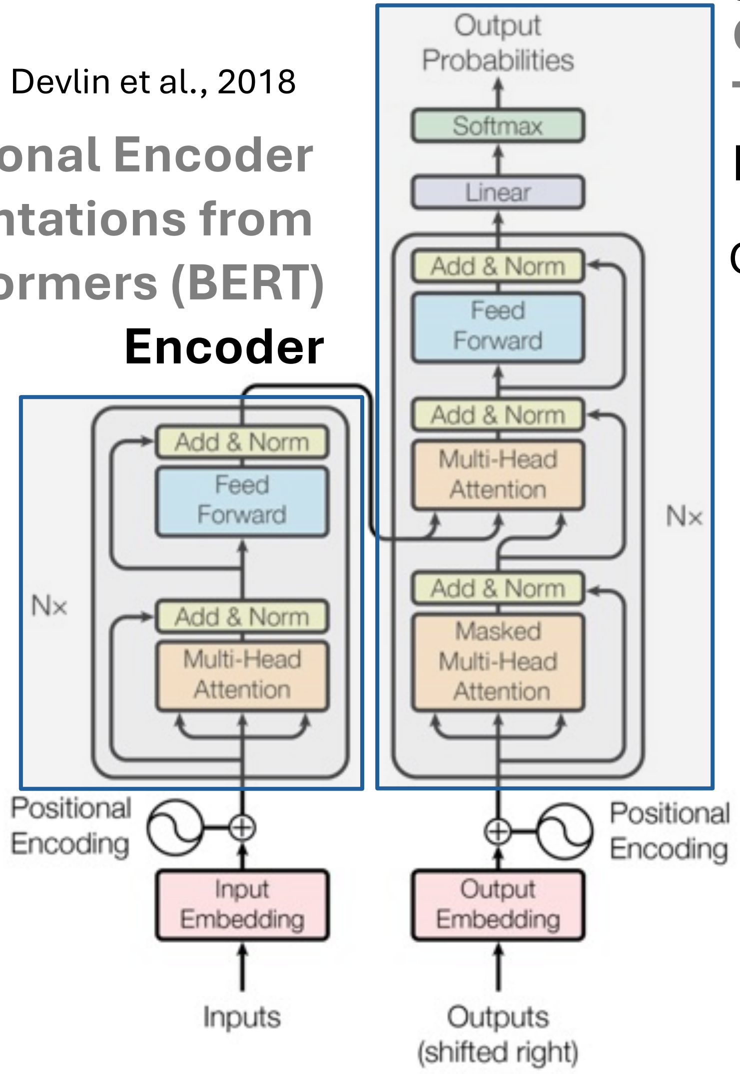
RoBERTa

Compact RoBERTa
Sanh et al., 2019

DistilRoBERTa

Climate-specific DistilRoBERTa
Webersinke et al., 2019

ClimateBERT



Transformer

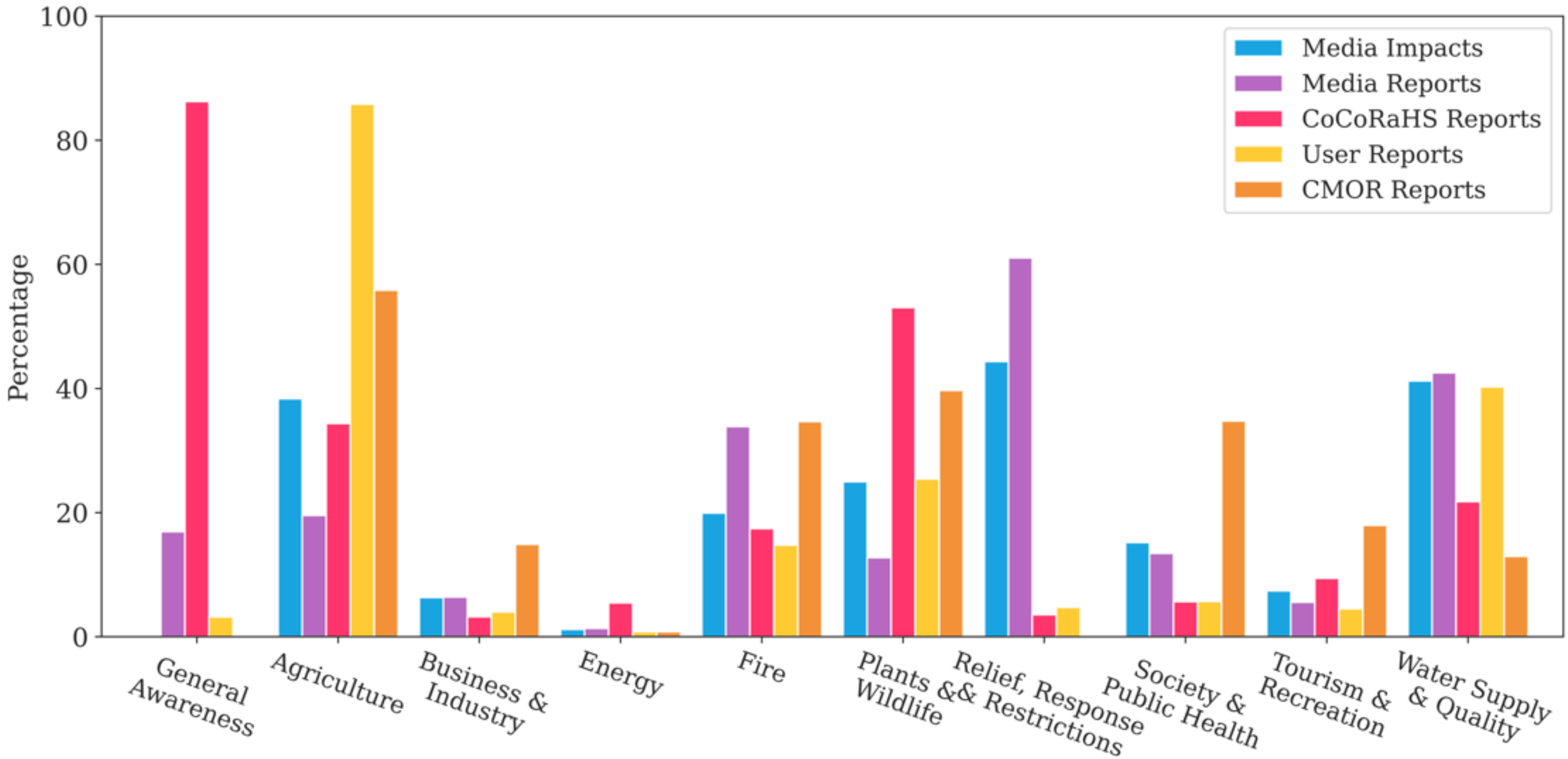
(Vaswani et al., 2018)

Benchmark Model:
Bag-of-Words **(BoW) +**
Term Frequency - Inverse Document Frequency **(TF-IDF) +**
Logistic Regression **(LR)**

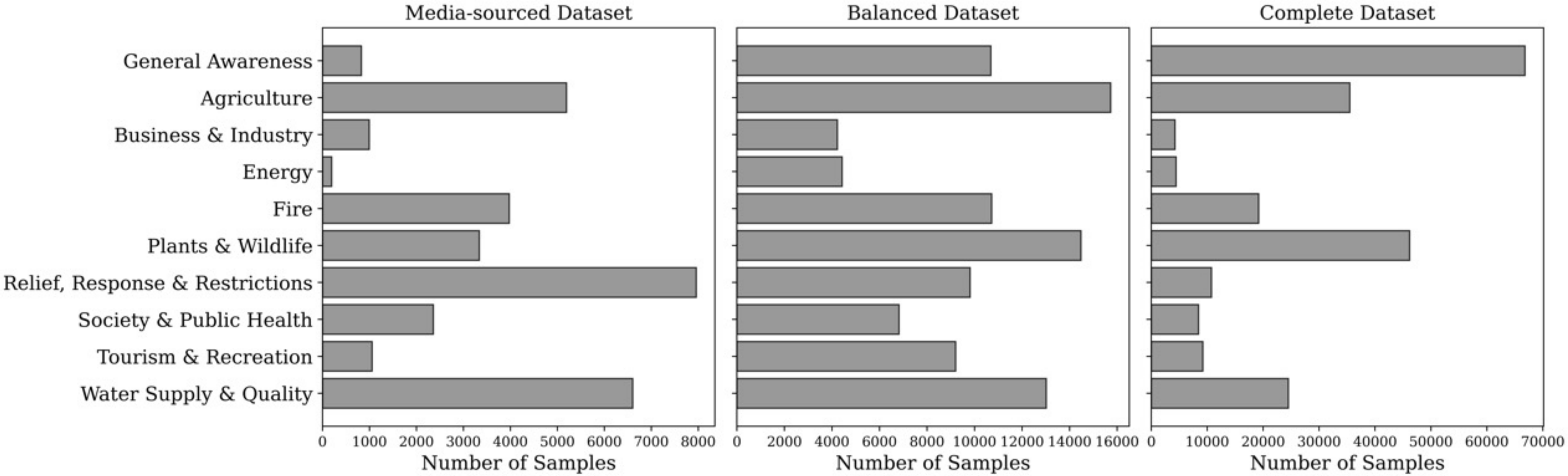
DIT and Social Media Datasets

	Period	Sample Count	Word Count in a Sample		
			Minimum	Median	Maximum
Media Impacts	2010-2023	11,878	16	69	462
Media Reports	2010-2023	5,684	22	399	478
CoCoRaHS Reports	2010-2023	76,428	5	36	488
User Reports	2010-2018	1,596	9	82	475
CMOR Reports	2019-2023	4,735	5	37	203
Total	-	100,321	5	44	488
Tweets	2020-2022	78,690	1	19	85

DIT Datasets



DIT Datasets for Fine-tuning BERT-like Models



Results – Performance Comparison on DIT Data

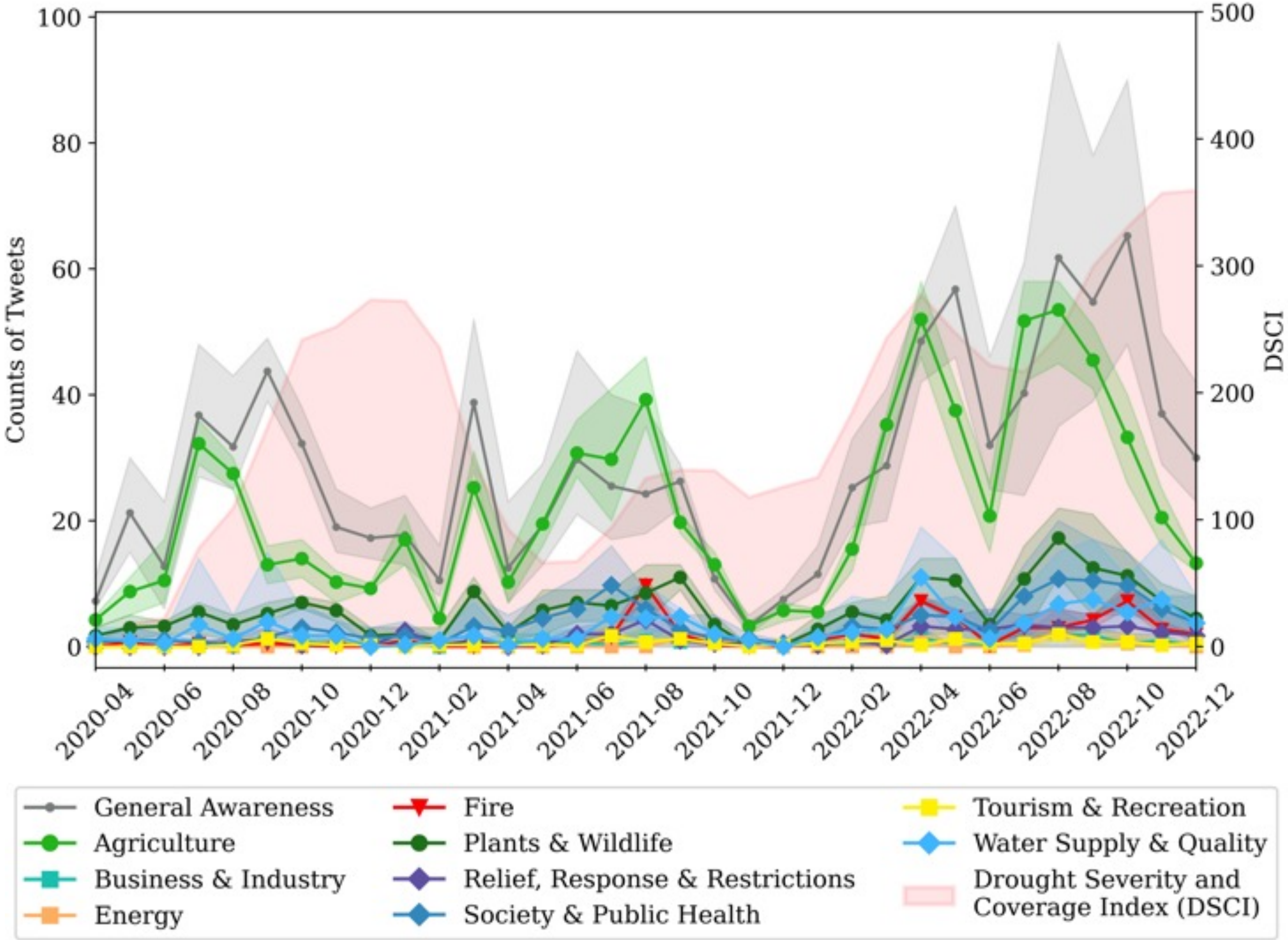
NLP Model	Datasets	Validation Dataset			Test Dataset		
		Exact Match Ratio	Hamming Loss	Weighted F1 Score	Exact Match Ratio	Hamming Loss	Weighted F1 Score
RoBERTa	**Media-sourced**	0.662	0.045	0.880	0.548	0.066	**0.840**
DistilRoBERTa	Media-sourced	0.662	0.046	0.874	0.552	0.064	0.833
ClimateBERT	Media-sourced	0.652	0.046	0.869	0.553	0.064	0.833
TF-IDF	-	0.465	0.081	0.684	0.372	0.107	**0.643**
RoBERTa	Balanced	0.483	0.101	0.820	0.549	0.065	0.828
DistilRoBERTa	Balanced	0.471	0.105	0.797	0.506	0.072	0.795
ClimateBERT	Balanced	0.477	0.106	0.801	0.530	0.069	0.817
TF-IDF	-	0.294	0.158	0.620	0.353	0.112	0.646
RoBERTa	Complete	0.458	0.096	0.765	0.522	0.070	0.806
DistilRoBERTa	Complete	0.446	0.099	0.758	0.491	0.076	0.788
ClimateBERT	Complete	0.431	0.101	0.749	0.467	0.081	0.761
TF-IDF	-	0.332	0.129	0.588	0.164	0.158	0.497
GPT-4	-	-	-	-	0.217	0.133	**0.778**

Results – Performance Comparison on Tweets

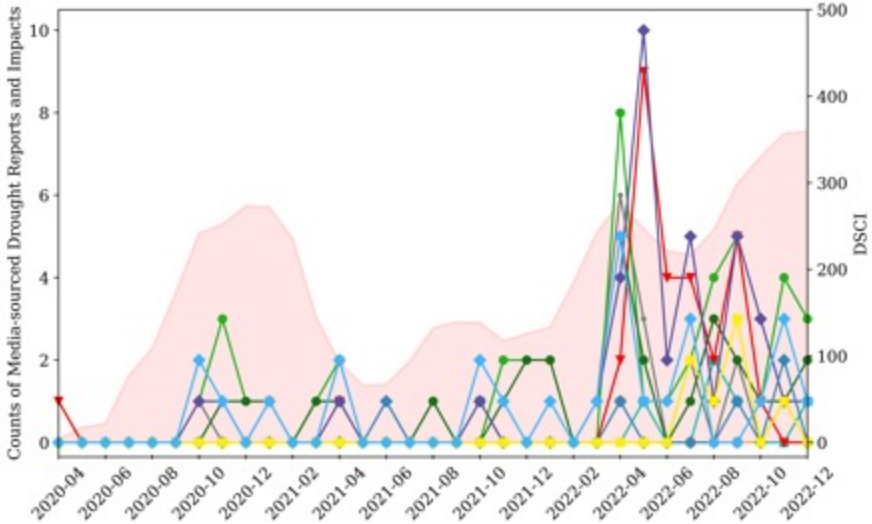
NLP Model	Fine-tuning Datasets	Prediction of Tweets		
		Exact Match Ratio	Hamming Loss	Weighted F1 Score
RoBERTa	Media-sourced	0.307	0.103	0.423
DistilRoBERTa	Media-sourced	0.251	0.120	0.359
ClimateBERT	Media-sourced	0.277	0.111	0.380
TF-IDF	-	0.341	0.088	0.236
RoBERTa	Balanced	0.191	0.197	0.407
DistilRoBERTa	Balanced	0.198	0.166	0.419
ClimateBERT	Balanced	0.156	0.207	0.443
TF-IDF	-	0.312	0.104	0.256
RoBERTa	Complete	0.390	0.090	0.574
DistilRoBERTa	Complete	0.335	0.102	0.557
ClimateBERT	**Complete**	0.390	0.093	**0.579**
TF-IDF	-	0.343	0.098	**0.368**
GPT-4	-	0.300	0.105	**0.608**

Results – Temporal Analysis in Nebraska

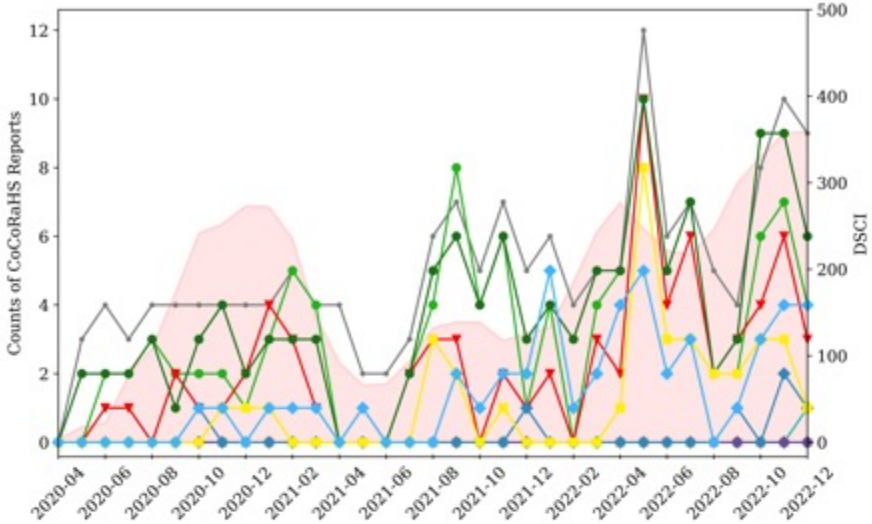
DL-based NLP Model Predicted Tweets



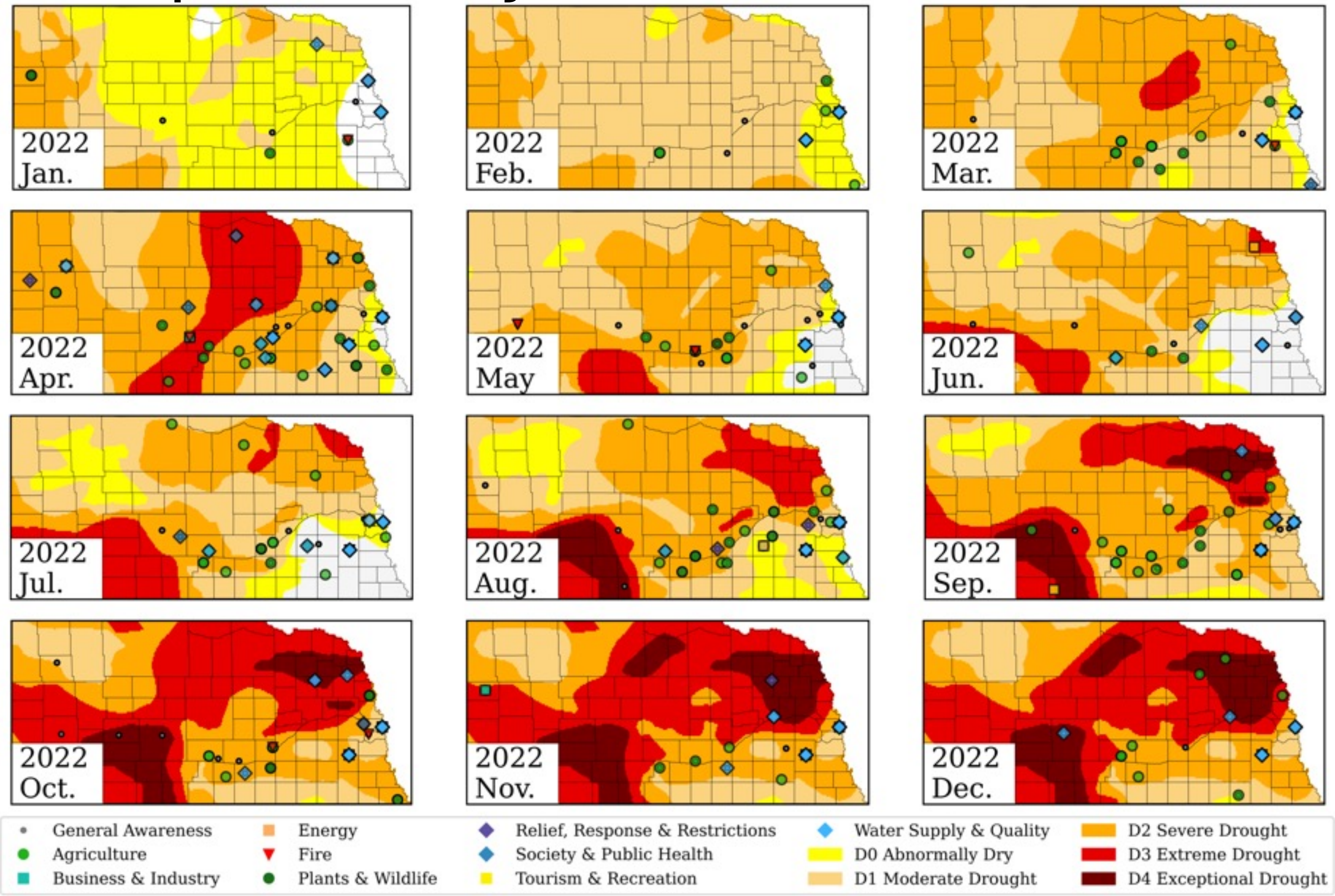
Media-sourced Drought Reports and Impacts



CoCoRaHS Reports



Results – Spatial Analysis in Nebraska



Discussion

Data
- Public awareness vs observed drought impacts on multiple sectors.
- Whether spatial patterns of the labeled tweets indicate bias or characteristics.

Methods
- Agentic prompting engineering.
- Retrieval Augmented Generation (RAG).
- Name Entity Recognition (NER) and Sentiment Analysis.

	BERT-like Models	GPT-4
Development	*Challenging* - Modifying and fine-tuning the pre-trained BERT-like models.	*Easy* - Prompt engineering through the API for GPT-4 from OpenAI.
Running Time	*Short* - Up to 49 hours for fine-tuning and 0.2 hours for predicting the tweet dataset.	*Long* - No fine-tuning process and 150 hours for predicting the tweet dataset.
Hardware Requirement	*High* - NVIDIA Tesla V100 32GB.	*Low* - No requirement.
Expense	*Low* - Primary cost is on fine-tuning, around $120. Cost of predicting tweets is less than $0.5.	*High* - Cost of the API from OpenAI is by word. It takes $730 for predicting tweets.
Scalability	*Low* - Multi-label classification task in the context of drought impacts.	*High* - Text generation with prompting that can be applied to other drought-related tasks.

Takeaways

Can we utilize the power of DL-based NLP models to assist users in categorizing drought impacts?

✓ We can use the RoBERTa model fine-tuned on the news-media-based dataset to accurately predict the label of drought impacts.

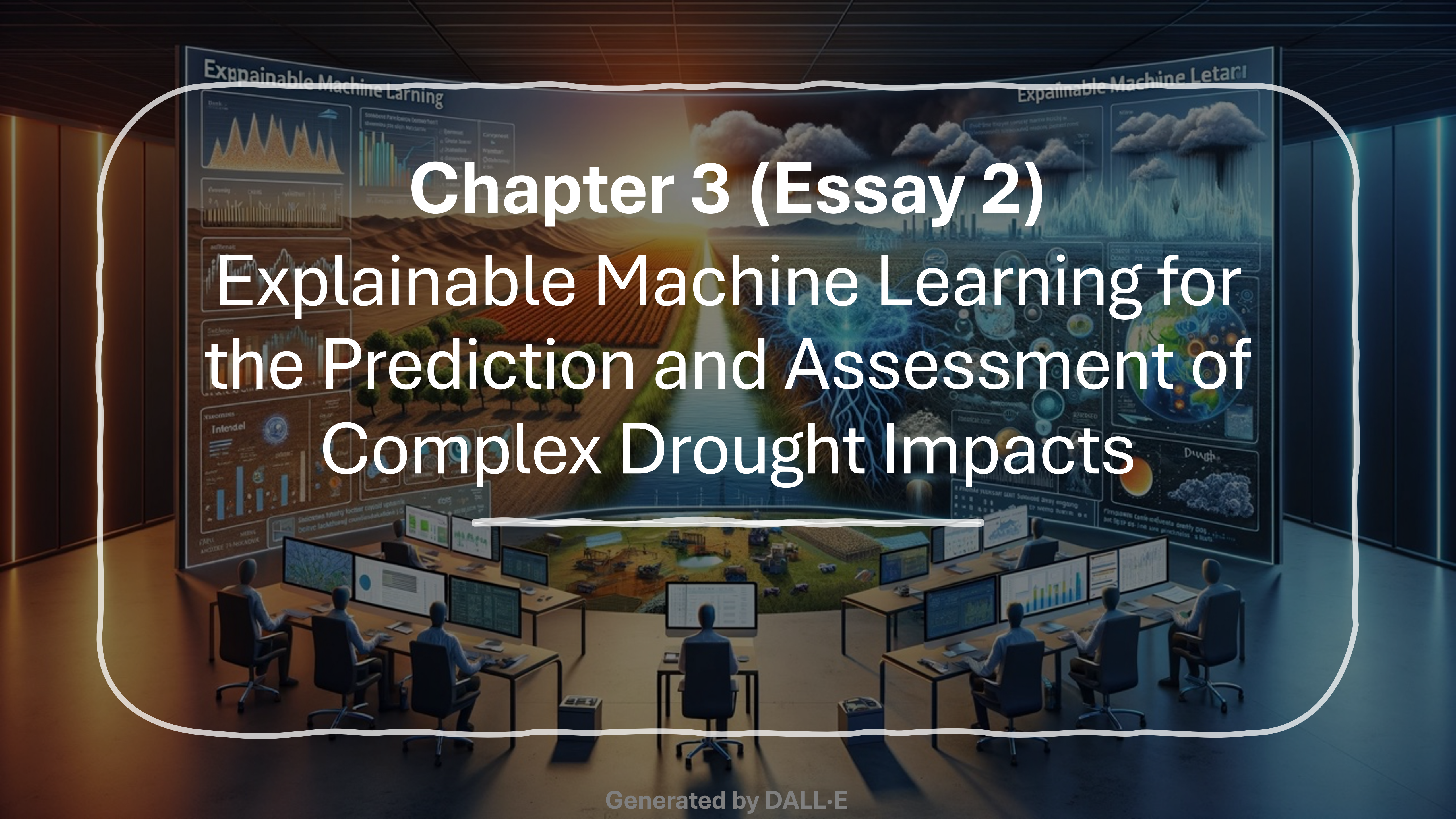
Can we apply the developed NLP models to help enlarge the drought impact database?

✓ We can use the ClimateBERT model fine-tuned on the complete dataset or the GPT-4 model with proper prompt engineering to predict the types of drought impacts indicated in the tweets.

Predicted labels of tweets can be used to provide **complementary information** to news media and volunteers' reports.

Spatial and temporal patterns reflected in the predicted labels of tweets in Nebraska reveal the **dynamic nature** of drought impacts.

The **relationships between textual data and drought indices** are beneficial for providing deep insights and improving understanding of drought impacts.



Chapter 3 (Essay 2)

Explainable Machine Learning for the Prediction and Assessment of Complex Drought Impacts

Generated by DALL·E

Brief Background and Research Justification

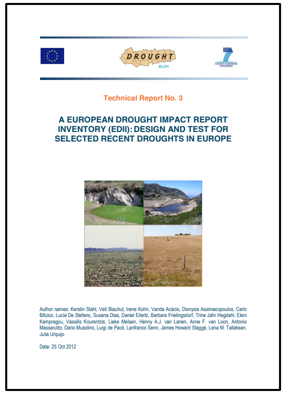
European Drought Impact Report Inventory (EDII)

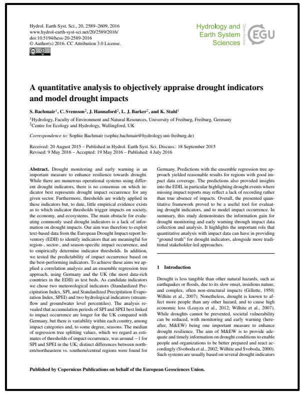
European Drought Impact Report Inventory (EDII) and European Drought Reference (EDR) database

EDCadmin | October 22, 2015 | 0 Comments

The European Drought Impact Report Inventory (EDII) and European Drought Reference (EDR) database are now available through a link under Resources (or by clicking here). The EDII allows the user to search reported drought impacts and submit new impact reports for Europe, while the EDR summarizes historical droughts for Europe and provides a tool to visualize SPI for any date (1958-2009).

Stahl et al., 2012, 2016, and many her group's studies.
Bachmair et al., 2015, 2016, 2017.





Drought Monitoring & Early Warning

Quantitative Impact Assessment & Prediction

Vulnerability & Risk Assessment

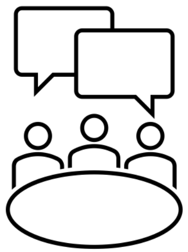
Policy & Planning

The relationship between drought indices and multifaceted drought impacts has not yet been quantitatively explored in the US.

Research Questions

Can we train and fine-tune a robust model through an ML pipeline setup to accurately predict imbalanced drought impacts?

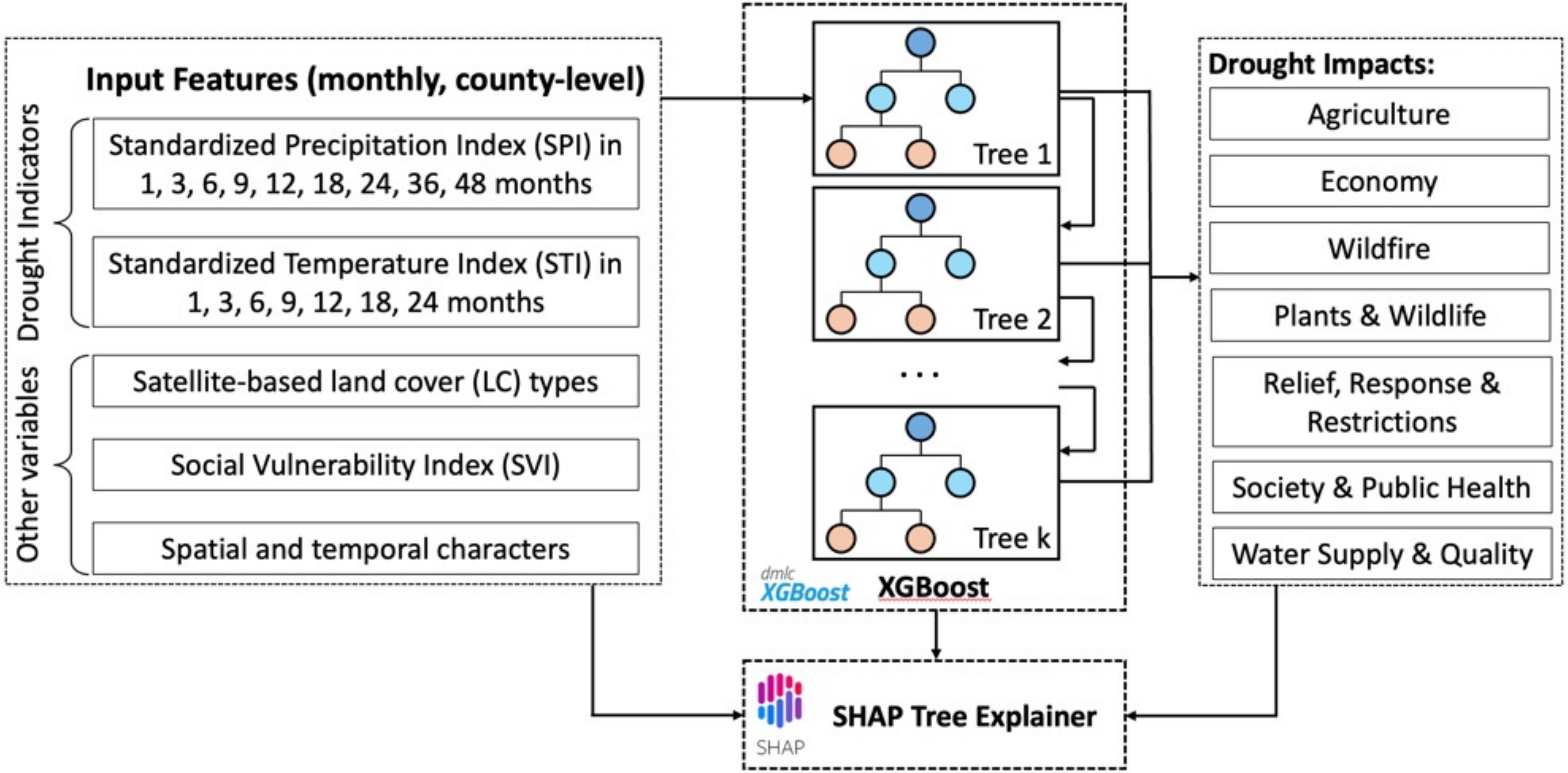
Is the explainability of the most accurate model beneficial for us to trust the model's prediction or further help us understand the relationship?



Research Objectives

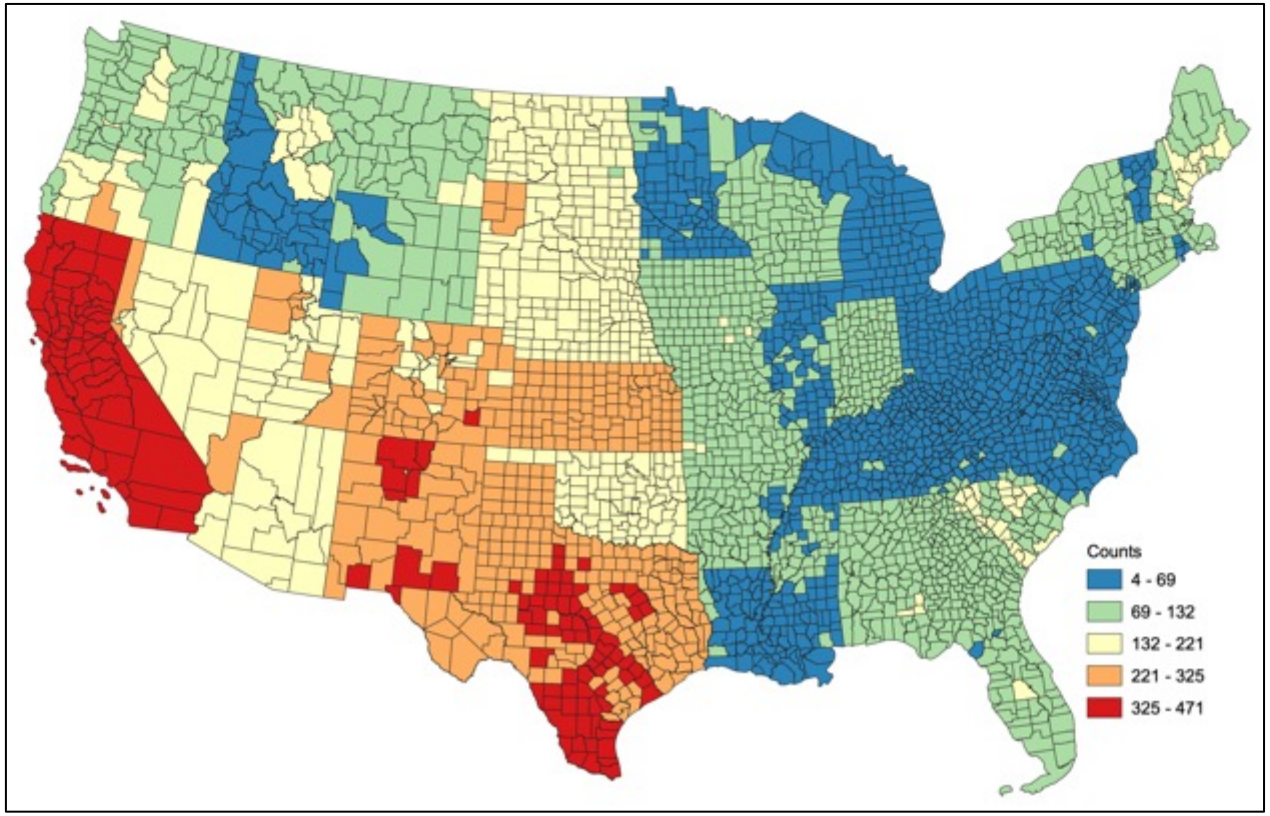
- Developing a robust explainable ML pipeline to predict drought impacts.
- Training and fine-tuning the models and comparing them with popular models.
- Exploring transfer learning of the models on predicting drought impacts.
- Applying an algorithm to explain the decision process of the ML models.
- Analyzing the interpreted relationships from the models in case studies.

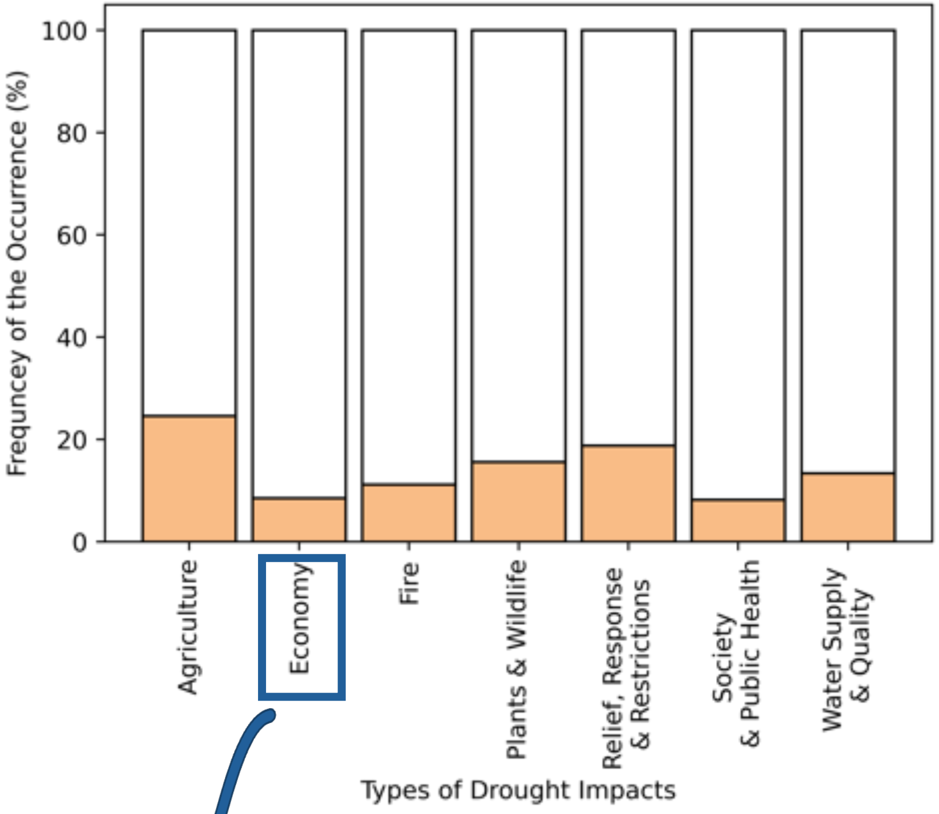
Explainable ML Pipeline

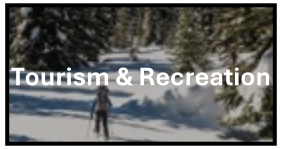


Multifaceted Drought Impacts in the CONUS

Counts of Drought Impacts based on NDMC DIT Datasets (2011-2020)







Results – Performance Comparison (CONUS-level)

Model	Performance Metrics	Dataset	Agriculture	Economy	Fire	Plants & Wildlife	Relief, Response & Restriction	Society & Public Health	Water Supply & Quality
XGBoost	F2 score	Validation	0.939	0.974	0.956	0.958	0.951	0.974	0.935
XGBoost	F2 score	Test	**0.914**	**0.881**	**0.876**	**0.899**	**0.889**	**0.868**	**0.856**
XGBoost	Recall	Validation	0.938	0.9	0.917	0.914	0.898	0.896	0.906
XGBoost	Recall	Test	**0.938**	**0.901**	**0.912**	**0.917**	0.899	**0.885**	**0.907**
RF	F2 score	Validation	0.79	0.904	0.84	0.835	0.79	0.938	0.842
RF	F2 score	Test	0.797	0.723	0.724	0.759	0.771	0.765	0.733
RF	Recall	Validation	0.906	0.846	0.904	0.89	0.917	0.827	0.861
RF	Recall	Test	0.905	0.841	0.905	0.889	**0.918**	0.827	0.799
LR	F2 score	Validation	0.567	0.804	0.795	0.673	0.689	0.814	0.737
LR	F2 score	Test	0.678	0.716	0.599	0.621	0.64	0.566	0.592
LR	Recall	Validation	0.89	0.804	0.765	0.843	0.802	0.773	0.758
LR	Recall	Test	0.89	0.527	0.763	0.844	0.803	0.773	0.767
SVM	F2 score	Validation	0.355	0.025	0.127	0.272	0.254	0.11	0.06
SVM	F2 score	Test	0.351	0.024	0.127	0.265	0.249	0.117	0.062
SVM	Recall	Validation	0.312	0.02	0.104	0.234	0.219	0.09	0.049
SVM	Recall	Test	0.307	0.019	0.105	0.228	0.214	0.097	0.051
OneR	F2 score	Validation	0.253	0.114	0.092	0.086	0.165	0.136	0.081
OneR	F2 score	Test	0.248	0.123	0.094	0.083	0.164	0.133	0.083
OneR	Recall	Validation	0.221	0.096	0.076	0.07	0.139	0.114	0.067
OneR	Recall	Test	0.213	0.103	0.078	0.068	0.138	0.111	0.069

Results – Case Study

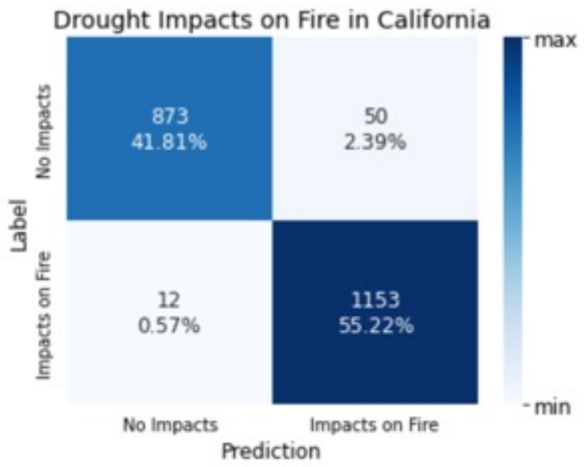
Drought Impacts Related to Wildfire in California

- The datasets are monthly at the county level from 2011 to 2020.
- The impacts are binarized into 0 and 1.
- The dimension of the input dataset: 6960 entries, 40 features.
- 70% of the data are used as the training and validation data, 30% are used as the test set.
- Nested cross-validation is used to fine-tune the model.

Performance metrics:

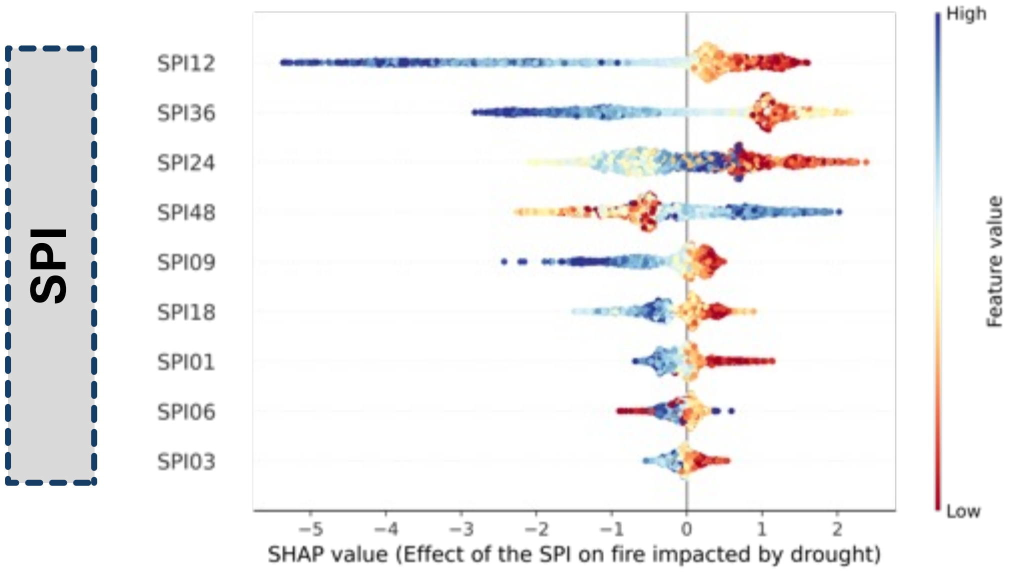
	Validation	Test
F2 score	0.97	0.98
Recall	0.99	0.99
Precision	0.91	0.96

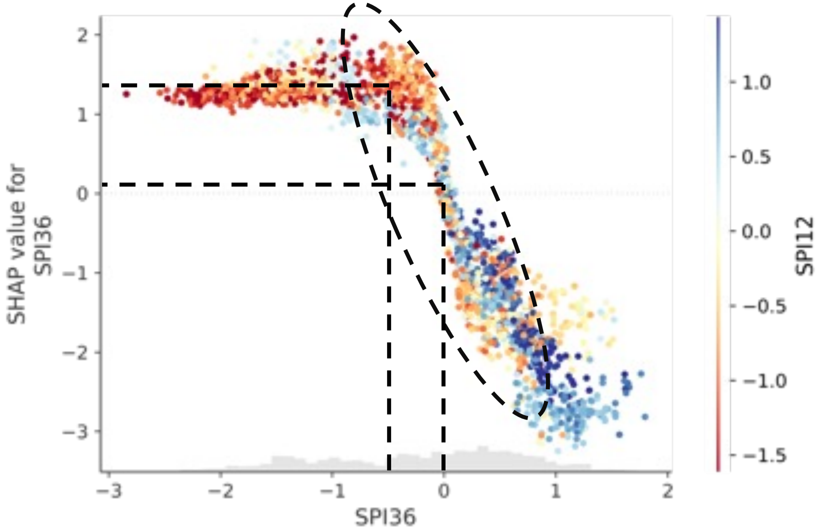
Confusion matrix (test dataset):

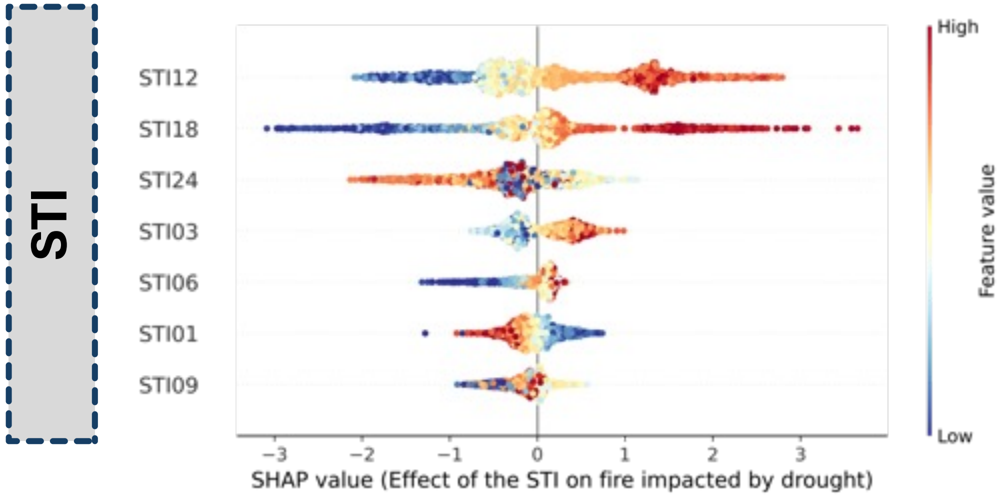


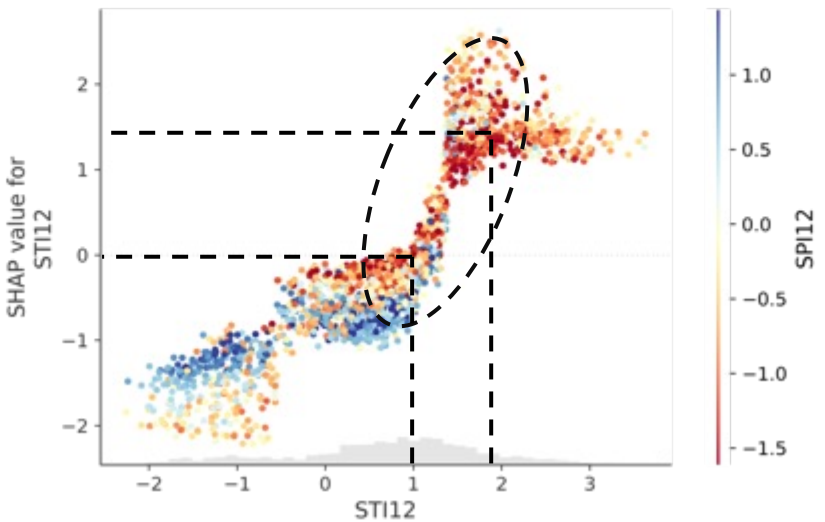
Results – Case Study

Drought Impacts Related to Wildfire in California





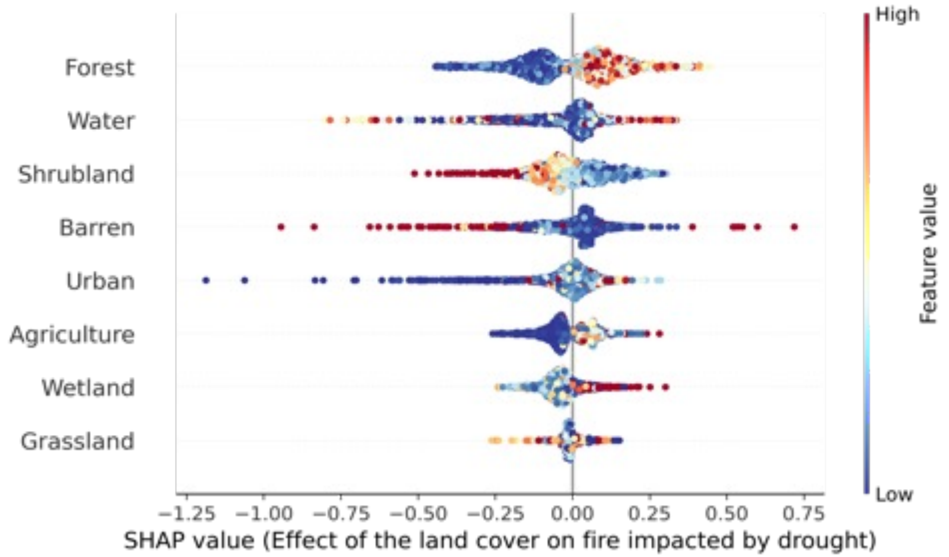


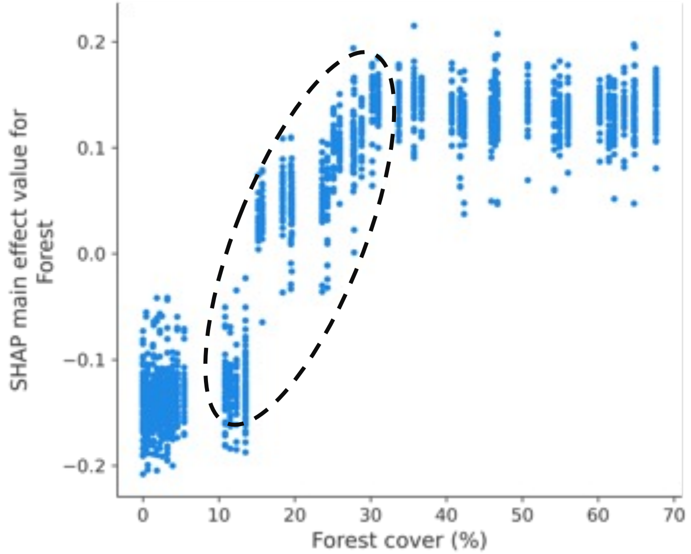


Results – Case Study

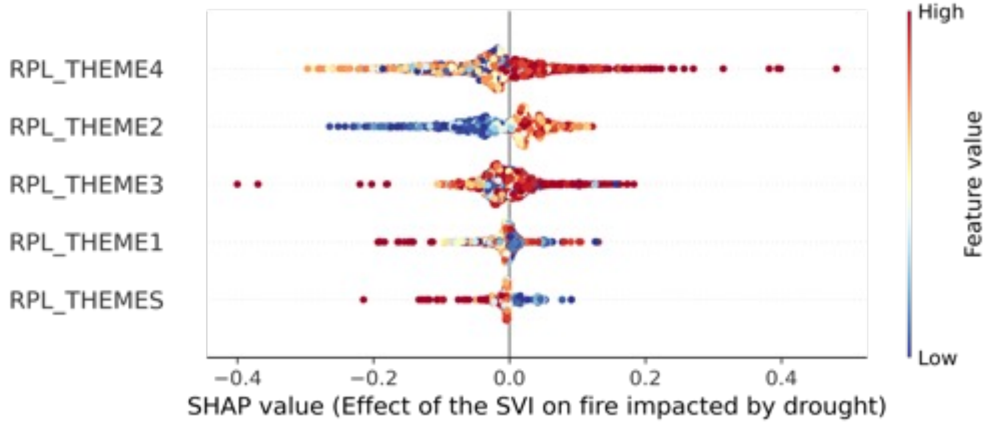
Drought Impacts Related to Wildfire in California

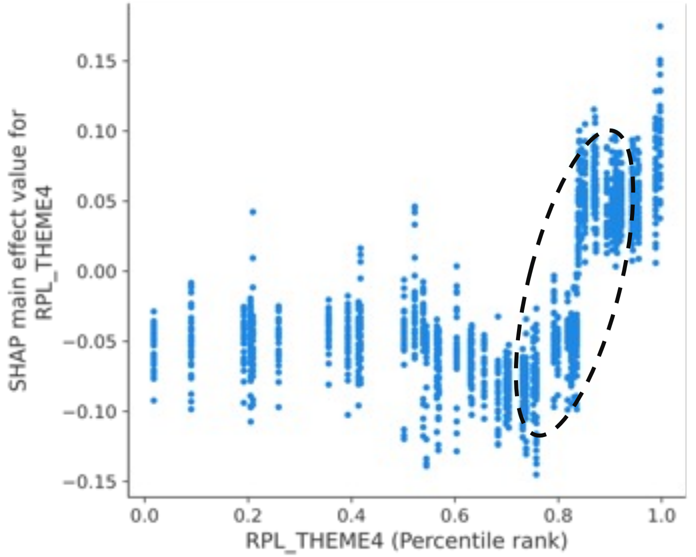
Land Cover





SVI





Discussion

What does trustworthy AI mean to the predictive studies of the impacts of climate/natural disasters?

Climate scientists

I would like to use ML models to help answer complex questions. **Can I trust the model?**

Computer scientists

I would like to use your domain expertise to validate the model explainability. **Can you tell me if the interpretation of the variables is correct?**

How to improve the interaction between the ML models/explainable tools (models) and the end users?

Takeaways

Can we train and fine-tune a robust model through an ML experimental setup to accurately predict imbalanced drought impacts?

✓ The fine-tuned XGBoost outperformed other models at the national level.

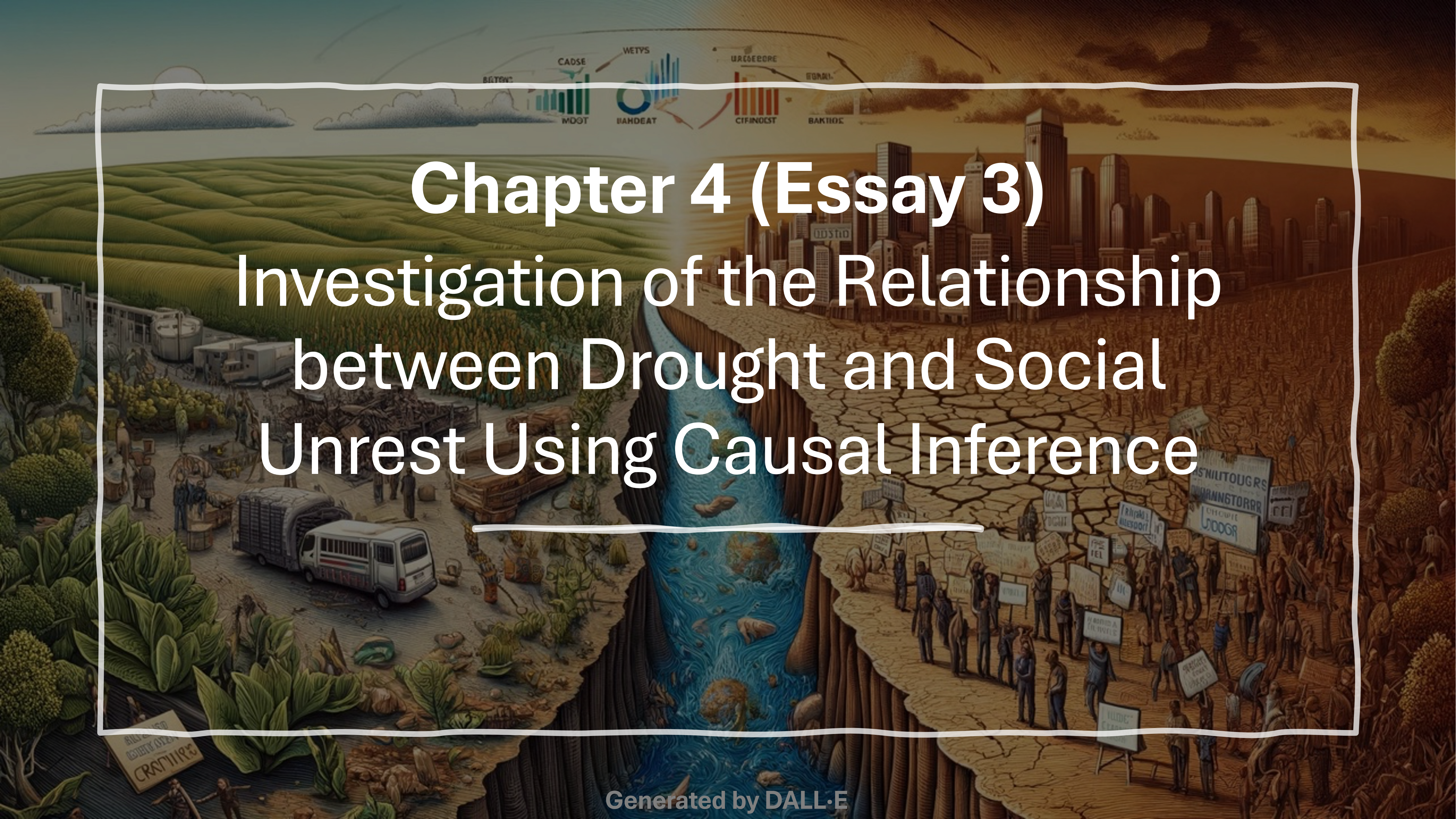
Is the explainability of the most accurate model beneficial for us to trust the model's prediction or further help us understand the relationship?

✓ At a certain level, Yes. There are explainable and non-linear patterns recognized by the models in the case study. However, there is also noticeable inconsistency in the relationships.

XGBoost does achieve a **very promising performance** on predicting **multifaceted** drought impacts.

In the California case study, **long-term precipitation and temperature anomalies** and **forest cover** are positively associated with the occurrence of drought impact on **wildfire**.

Further studies focus on **a certain type of drought impact** with an identified **impact pathway** might help clarify the relationship between drought indices and the drought impact.



Chapter 4 (Essay 3)

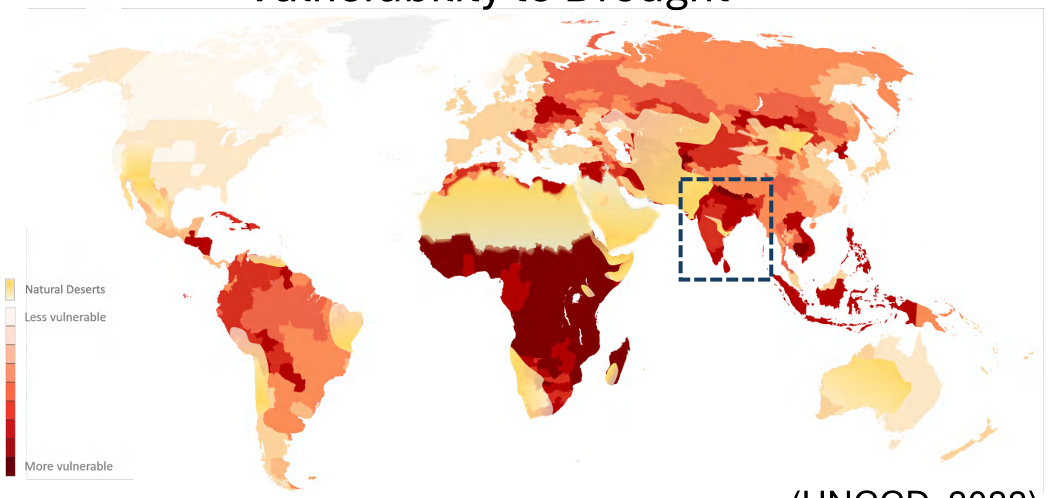
Investigation of the Relationship between Drought and Social Unrest Using Causal Inference

Generated by DALL·E

Brief Background and Research Justification

Countries Affected by Drought in 2022

Vulnerability to Drought



(UNCCD, 2022)

Increasing frequency of
protest events

?

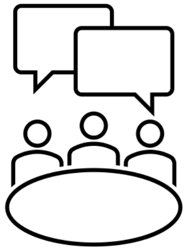
Timeline

Drought events
For example, deficient rainfall
in the monsoon season

Research Questions

Can we build a potential outcomes framework to estimate and control the effect of covariates?

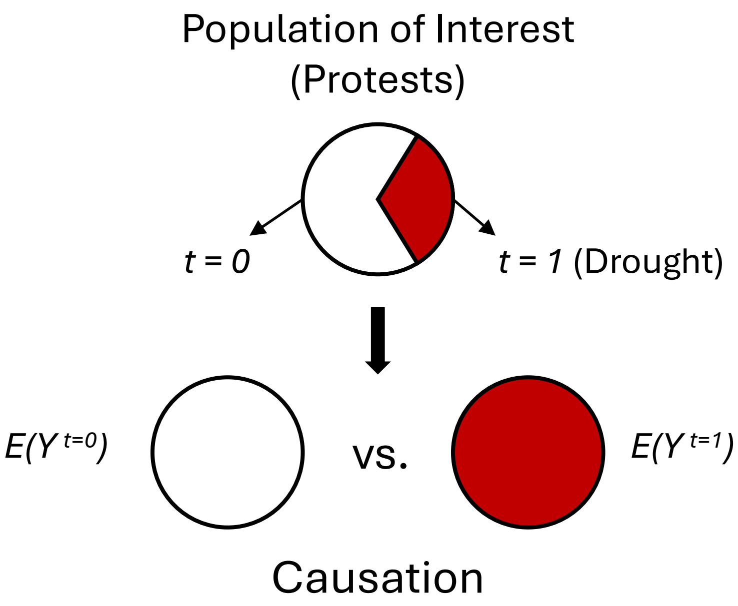
Is there any evidence for a causal relationship between extreme drought and increasing social unrest?

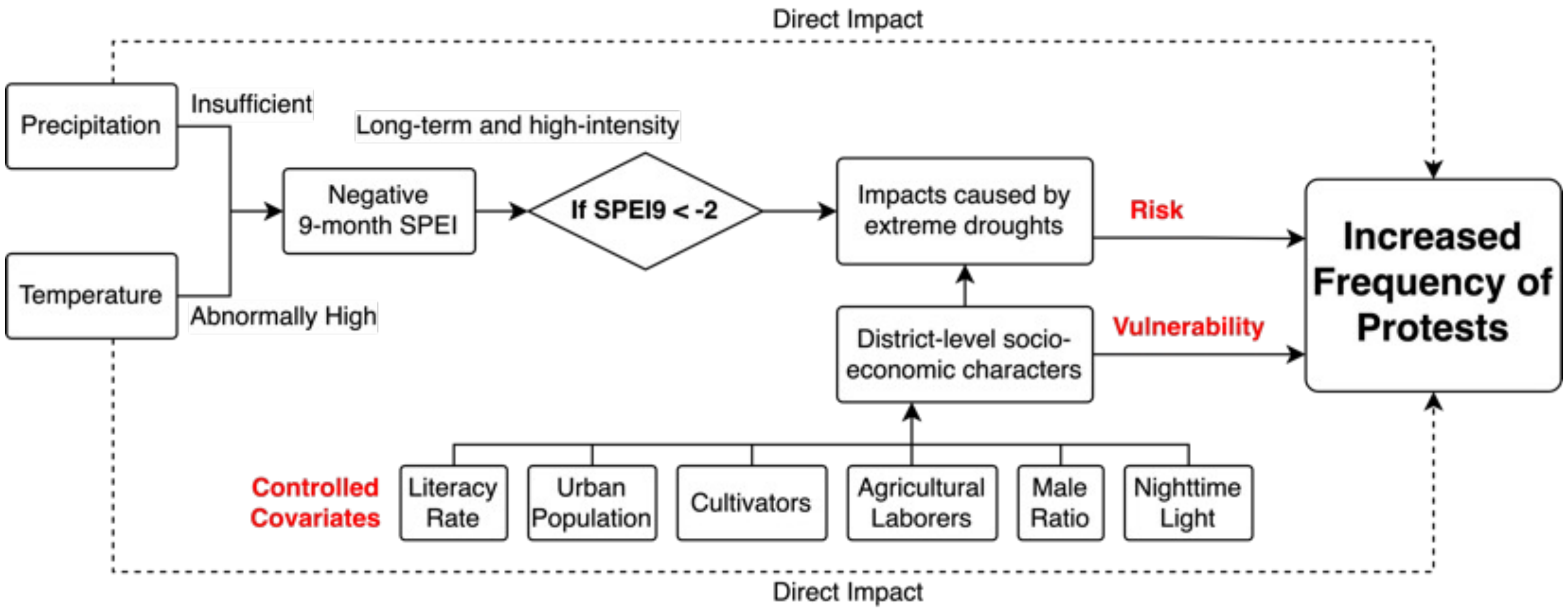


Research Objectives

- Data exploratory analysis for spatial and temporal patterns of drought and social unrest in India.
- Developing a potential outcomes framework (Rubin Causal Model) to estimate the average causal (treatment) effect of extreme drought on the increasing frequency of protests.
- Sensitivity analysis of the estimated average causal effect.

Causality and Directed Acyclic Graph (DAG)



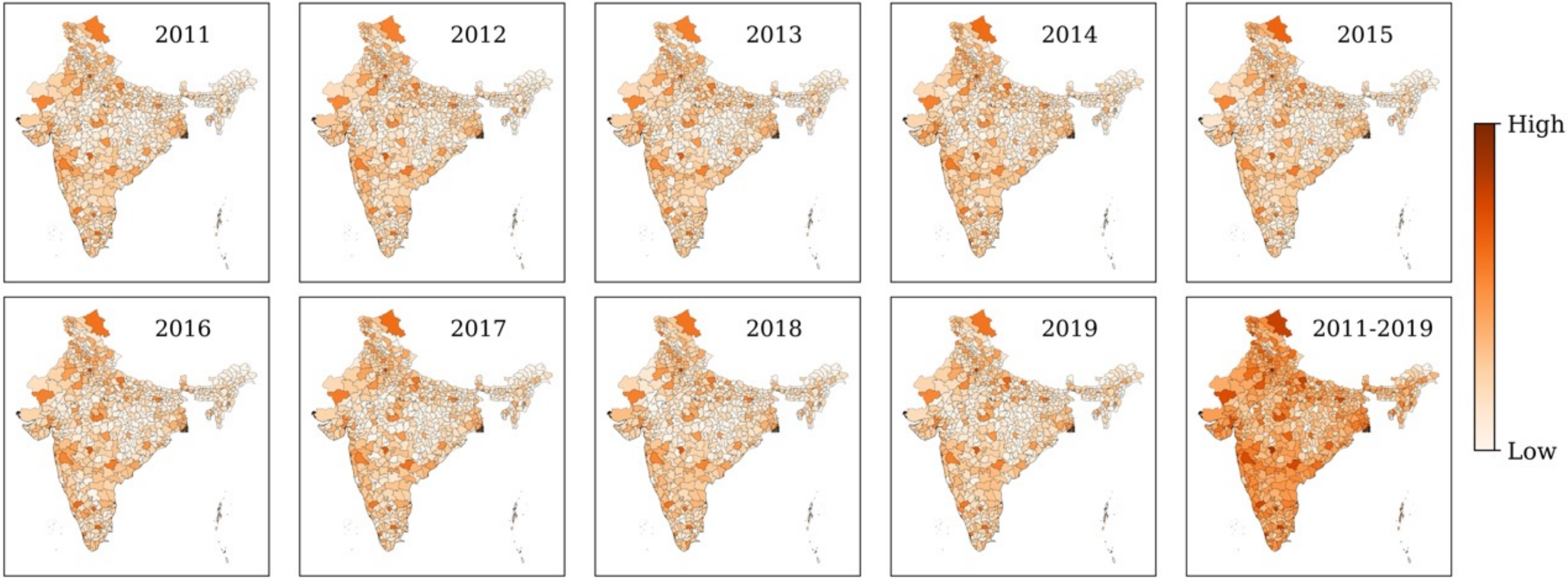


Data Description

	Dataset	Data Type	Spatial Resolution	Temporal Resolution	Period
Social Unrest	Protests	Polygon	District	Monthly	2010-2018
Drought	SPEI	Polygon	District	Monthly	2010-2019
Climate	Air Temperature	Raster	0.5 degree	Monthly	2010-2019
	Precipitation	Raster	0.05 degree	Monthly	2010-2019
	Soil Moisture	Raster	0.25 degree	3-day	2010-2019
Satellite	Nightlight	Raster	0.008 degree	Annual	2013
Census	Agricultural Labors	Polygon	District	-	2011
	Cultivators	Polygon	District	-	2011
	Literacy	Polygon	District	-	2011
	Male	Polygon	District	-	2011
	Urban Population	Polygon	District	-	2011

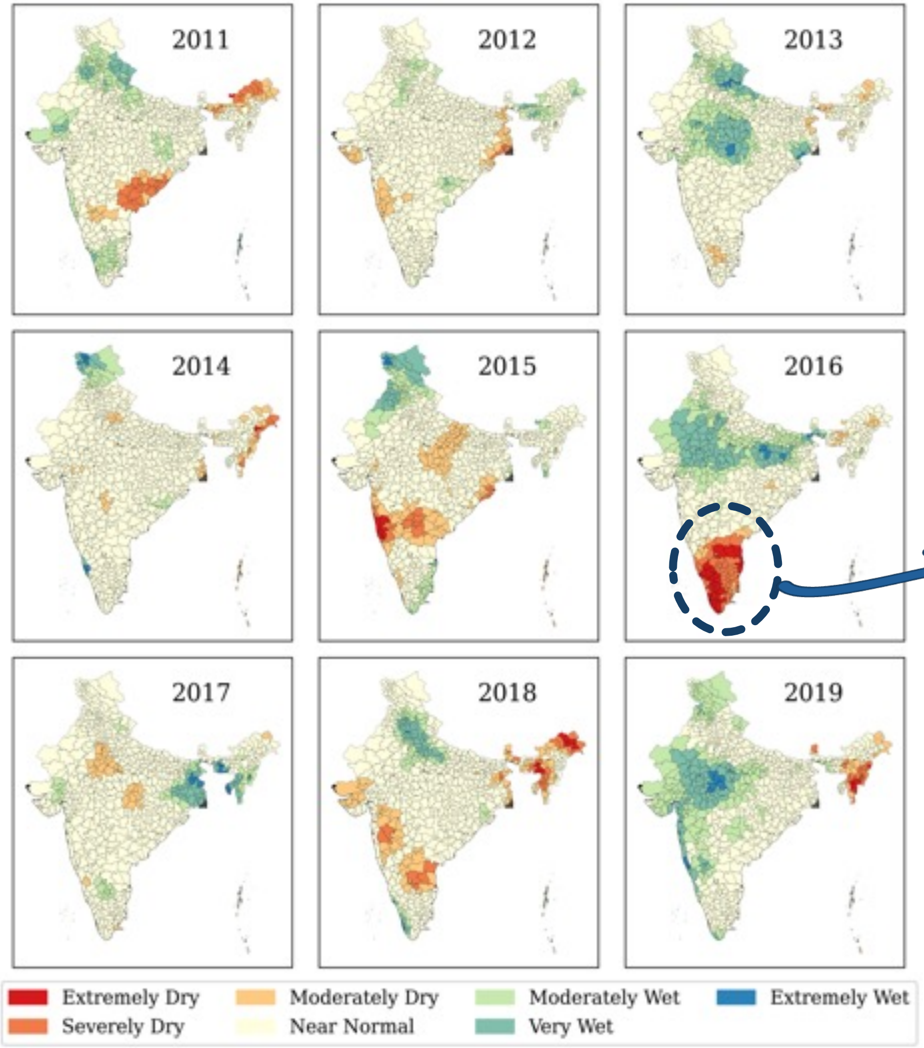
Data Exploratory Analysis

Annual count of protests in India by district from 2011 to 2019

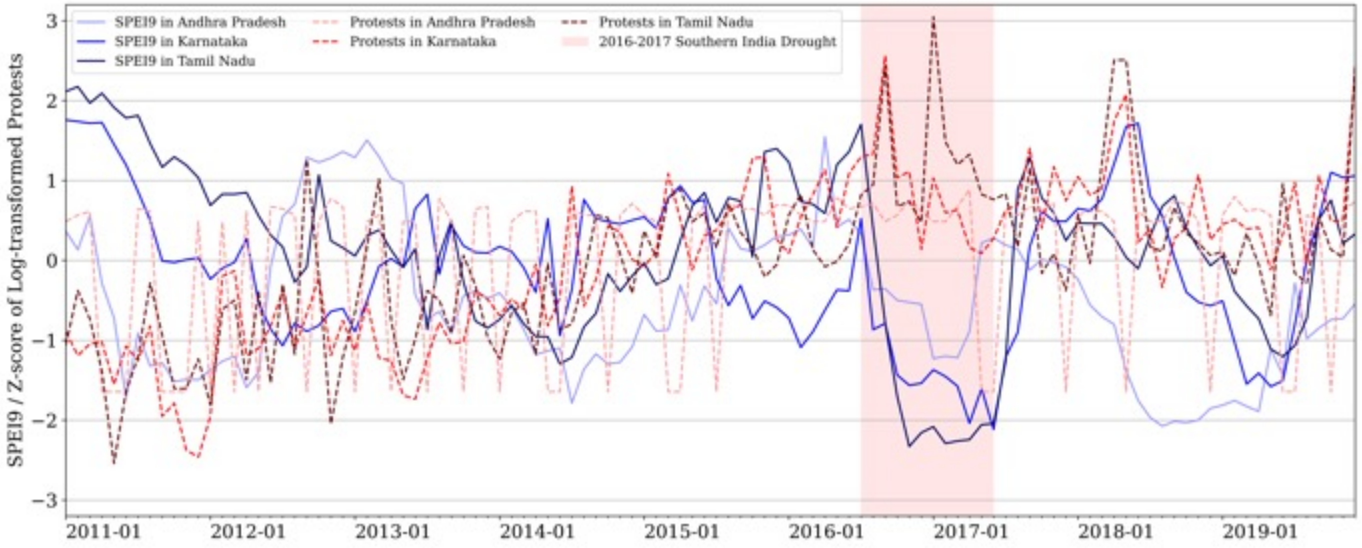


Data Exploratory Analysis

12-month SPEI in December from 2011 to 2019



2016-2017 South India Drought (**the worst in 150 years**)



Results – Propensity Score Model

Observations: 66096
Dependent Variable: drought (SPEI9 < -2)

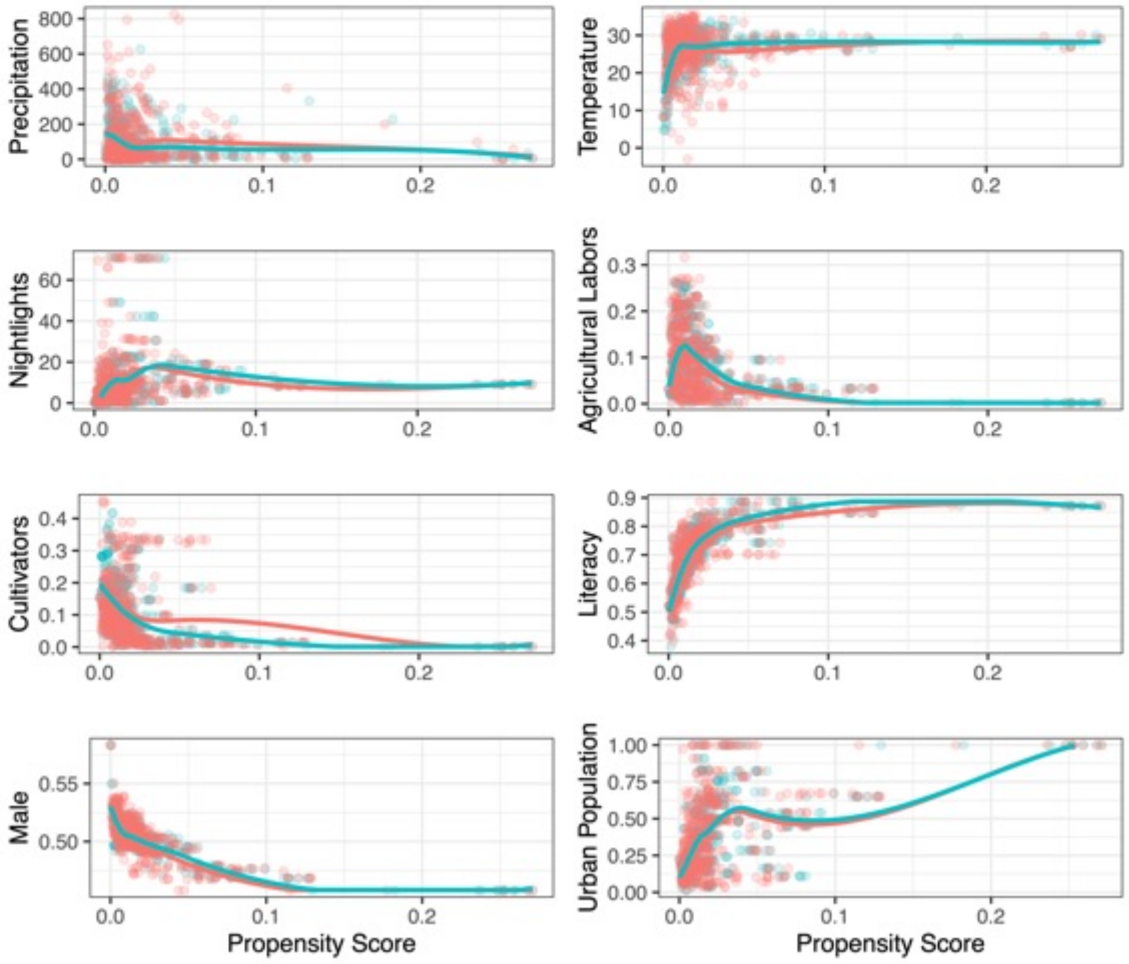
	Estimate	Std. Error	z value	p value	
(Intercept)	8.824	1.755	5.027	4.97E-07	***
Precipitation	-2.38E-03	3.39E-04	-7.027	2.11E-12	***
Temperature	0.005	9.55E-03	5.441	5.30E-08	***
Nightlights	-7.71E-03	5.58E-03	-1.381	0.167	
Agricultural Labors	1.829	0.083	-2.2	0.028	*
Cultivators	2.39	0.672	3.555	3.78E-03	***
Literacy	4.457	0.618	7.21	5.61E-13	***
Male	-35.44	2.963	-11.962	< 2e-16	***
Urban Population	1.101	0.281	3.923	8.75E-05	***

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Results – Propensity Score Matching

mean(SD)	**No Drought** (*t = 0*)	**Drought** (*t = 1*)	**SMD**
Precipitation	96.27 (146.66)	91.82 (102.11)	0.04
Temperature	25.82 (5.35)	25.80 (4.92)	0.01
Nightlights	10.18 (11.40)	9.60 (11.28)	0.05
Agricultural Labors	0.09 (0.07)	0.09 (0.07)	0.01
Cultivators	0.12 (0.10)	0.11 (0.09)	0.04
Literacy	0.71 (0.10)	0.70 (0.10)	0.07
Male	0.50 (0.02)	0.50 (0.02)	0.05
Urban Population	0.37 (0.25)	0.36 (0.24)	0.02

Number of matched samples: 535



Results – Estimation of Average Treatment Effect

Est. Coef. (Std. Err.)	IPTW Model (Full data)	Reduced Model (Matched data)	Full Model (Matched data)
(Intercept)	-0.033 (0.003) ***	-0.147 (0.042) ***	0.787 (1.263)
Drought	0.268 (-0.096) **	0.272 (0.060) ***	0.270 (0.006) ***
Precipitation	-	-	8.787e-5 (2.462e-4)
Temperature	-	-	-0.012 (0.008)
Nightlights	-	-	0.003 (0.004)
Agricultural	-	-	0.957 (0.593)
Cultivators	-	-	0.352 (0.458)
Literacy	-	-	-0.614 (0.426)
Male	-	-	-0.787 (2.14)
Urban Population	-	-	0.092 (0.196)

*Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Results – Rosenbaum Sensitivity Test

Gamma	Lower bound	Upper bound
1	0.026	0.026
1.1	0.002	0.157
1.2	0.001	0.438
1.3	0	0.735
1.4	0	0.912
1.5	0	0.979
1.6	0	0.996
1.7	0	0.999
1.8	0	0.999
1.9	0	1
2	0	1

✓ Based on the estimated propensity scores, there are statistically significant average treatment effect observed from the linear regression model.

?? The estimated average treatment effect is sensitive and not robust to unmeasured confounding variables (covariates).

Discussion

How can we improve the estimation of causal relationship between severe drought and protests?

Data

- Time-invariant socioeconomic characters.
- Insufficient direct observations of drought impacts on sectors, such as water supply and quality and crop failure.
- Assumption that severe drought events are equivalent to drought impacts could be various over the space and time.

Methods

- Developing other potential DAGs.
- Applying non-linear and complex models for estimating the causal relationship after propensity score matching or weighting.
- Experiencing other types of causal ML models/frameworks, such as Granger causality test for time-series data or Difference in Differences for mixed-type covariates.

Takeaways

Can we build a potential outcomes framework to estimate and control the effect of covariates?

- ✓ Based on the developed DAG and corresponding propensity score model, we could match the treated and control groups based on the covariates.

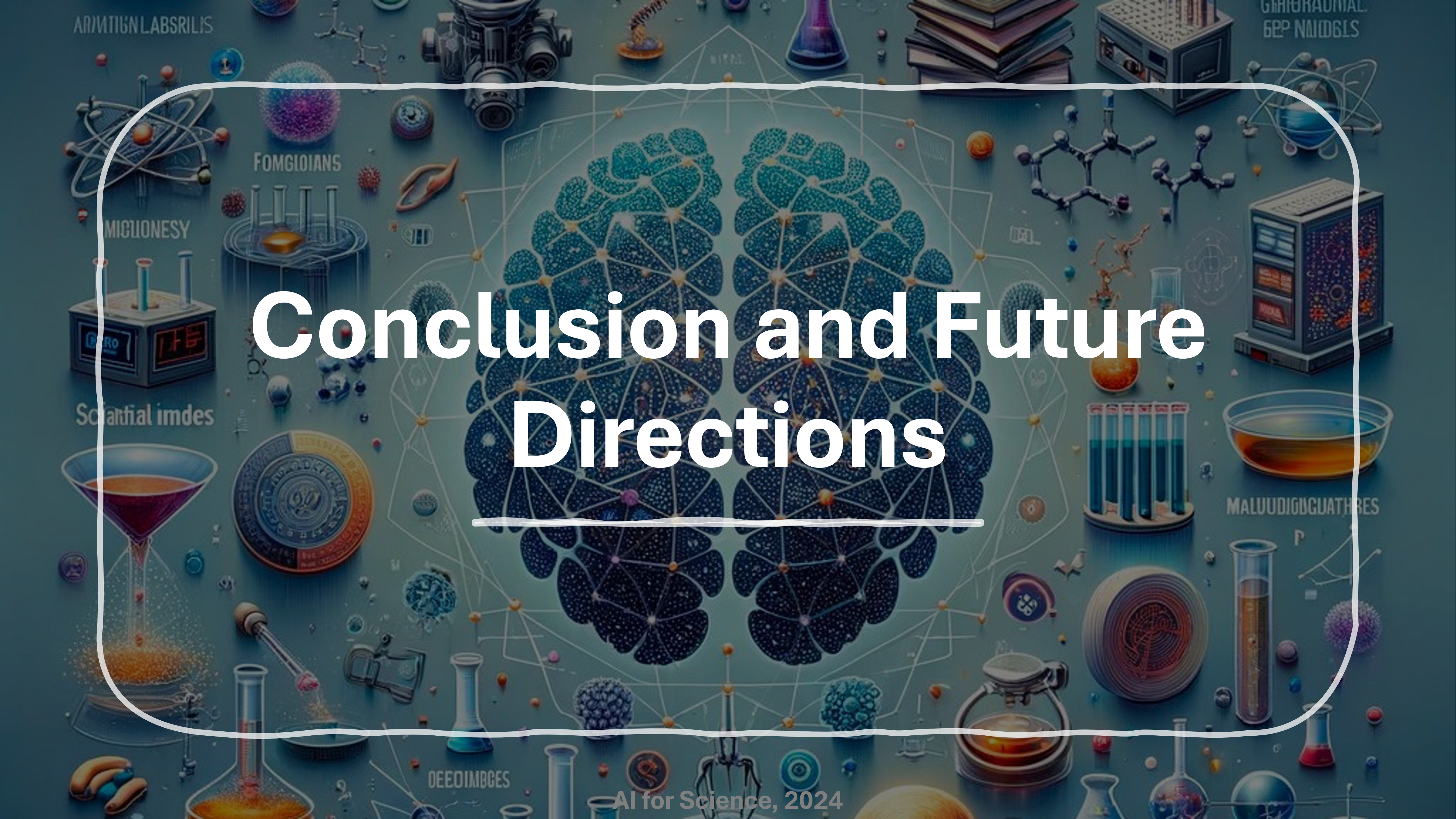
Is there any evidence for a causal relationship between extreme drought and increasing social unrest?

- ✓ The matched and weighted regression models all revealed a statistically significant average treatment effect of extreme drought on the increasing frequency of protests in India. However, the causal relationship is sensitive to unmeasured confounding variables.

Through the Rubin Causal Model, our study shows evidence of **a potential causal relationship between drought and protests in India.**

Time-invariant covariates describing the societal and economic characters can **reduce the con-founding effects** at the spatial level.

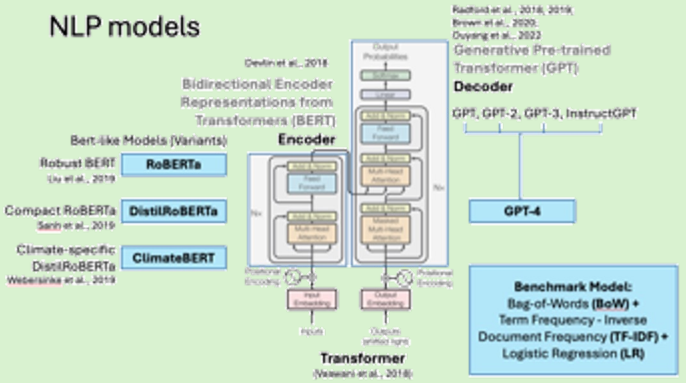
Further investigations are encouraged to focus on **improving the quality and diversity of the covariates** and building a more **robust causal model** with advanced ML algorithms.



Conclusion and Future Directions

Brief Recap of the Three Research Chapters

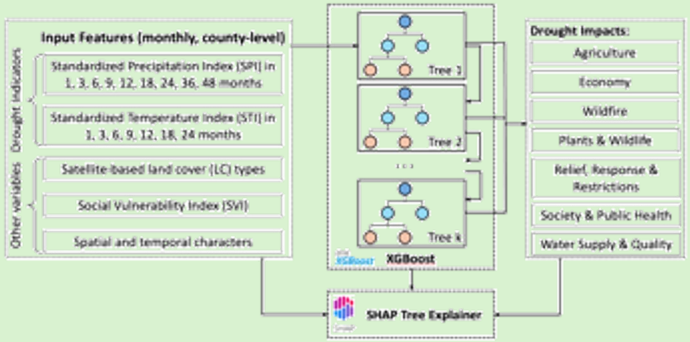
NLP models



Research Goal:
Improve the *datasets* of multi-dimensional drought impacts.

Significance and Contribution:
DL-based NLP models *improved and simplified the categorization process* significantly and *confirmed and demonstrated tweets can be used to monitor drought impacts* in the US.

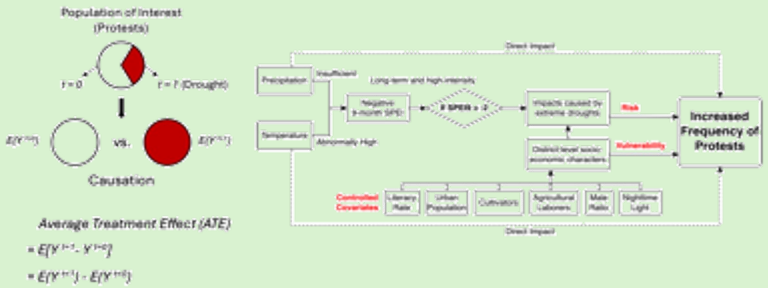
Explainable ML Pipeline



Research Goal:
Identify *relationships* between drought indices and impacts.

Significance and Contribution:
The explainable ML pipeline, XGBoost and SHAP, showed the *best predictive performance* and uncovered *the non-linear and complex relationship* of drought impacts on fire in California.

Causality and Directed Acyclic Graph (DAG)



Research Goal:
Investigate *causality* between extreme drought and protests.

Significance and Contribution:
Rubin Causal Model illustrated a statistically significant but not robust causal relationship between extreme drought and the increasing frequency of protests in India.

	Gaps	AI Methods	Outcomes
Essay 1	Insufficient and constrained quantitative datasets to observe and describe drought impacts beyond agricultural and ecological sectors.	NLP & DL	Assist data collection and expand a new data source
Essay 2	Understudied links between drought indices and drought impacts on multi-dimensional sectors.	Explainable ML	Interpret relationships between drought indices and multifaceted impacts.
Essay 3	Particularly, lack of studies on drought impacts on socioeconomic sectors.	Causal ML	Identify a causal relationship between drought and the frequency of protests.

Outcomes from this Dissertation

Peer-reviewed Articles

Zhang, B., Salem, F. K. A., Hayes, M. J., Smith, K. H., Tadesse, T., & Wardlow, B. D. (2023). Explainable machine learning for the prediction and assessment of complex drought impacts. *Science of The Total Environment*, 165509.

Zhang, B., Schilder, F., Smith, K., Hayes, M., Harms, S., & Tadesse, T. (2021). TweetDrought: A deep-learning drought impacts recognizer based on Twitter data. *ICML Workshop on Tackling Climate Change with Machine Learning*. [Workshop Preprint]

Zhang, B., Abu Salem, K. F., Hayes, M., & Tadesse, T. (2020). Quantitative assessment of drought impacts using XGBoost based on the Drought Impact Reporter. *NeurIPS Workshop on Tackling Climate Change with Machine Learning*. [Workshop Preprint]

Article Under Review

Zhang, B., Smith, K. H., Hayes, M. J., Tadesse, T., Schilder, F., Salem, F. K. A., & Samal, A. Tracking Drought Impacts from Texts: Towards AI-Assisted Drought Impact Detection. *Bulletin of the American Meteorological Society.*

Article in Progress

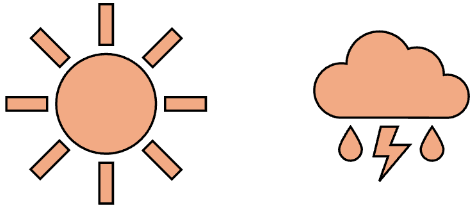
Zhang, B., Hayes, M. J., Shi, X., Werum, Tadesse, T., R, Guan, Y., Zhou, Y., Samal, A., A. Tracking Drought Impacts from Texts: Towards AI-Assisted Drought Impact Detection. Investigation of the Relationship of Drought and Social Unrest Using Causal Inference.

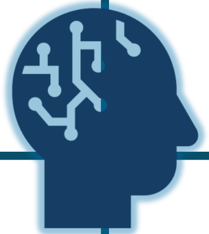
Broad Impacts from this Dissertation

AI for Science and for Good

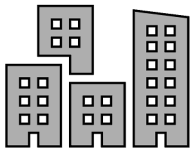
Water Resource Management & Sustainability Development

Natural Disaster Monitoring & Impact Assessment

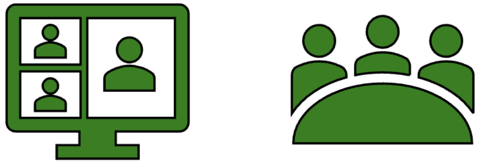




Proactive Planning for Climate Resilience



Decision Supports for Climate Change Adaptation



Essay 1
Assist data collection and expand a new data source

Essay 2
Interpret relationships between drought indices and multifaceted impacts.

Essay 3
Identify a causal relationship between drought and the frequency of protests.

Trustworthy

Ethical

Human-centered

And many more …

"I want an AI-powered society because I see so many ways that AI can make human life better"

"Responsible AI is important, and AI has risks"

- Andrew NG

Acknowledgement

University of Nebraska:

National Drought Mitigation Center

Dr. Kelly Helm Smith, Dr. Mark Svoboda

University of Nebraska Medical Center

Dr. Jesse Bell

Department of Sociology

Dr. Regina Werum

Daugherty Water For Food Global Institute

American University of Beirut

Dr. Fatima K. Abu Salem

Thomson Reuters Labs

Dr. Frank Schilder

Ph.D. students I have been working with …

My parents, wife, other family members, and friends …

Thank you!

ANY QUESTIONS OR COMMENTS ARE WELCOME!

Stay connected!