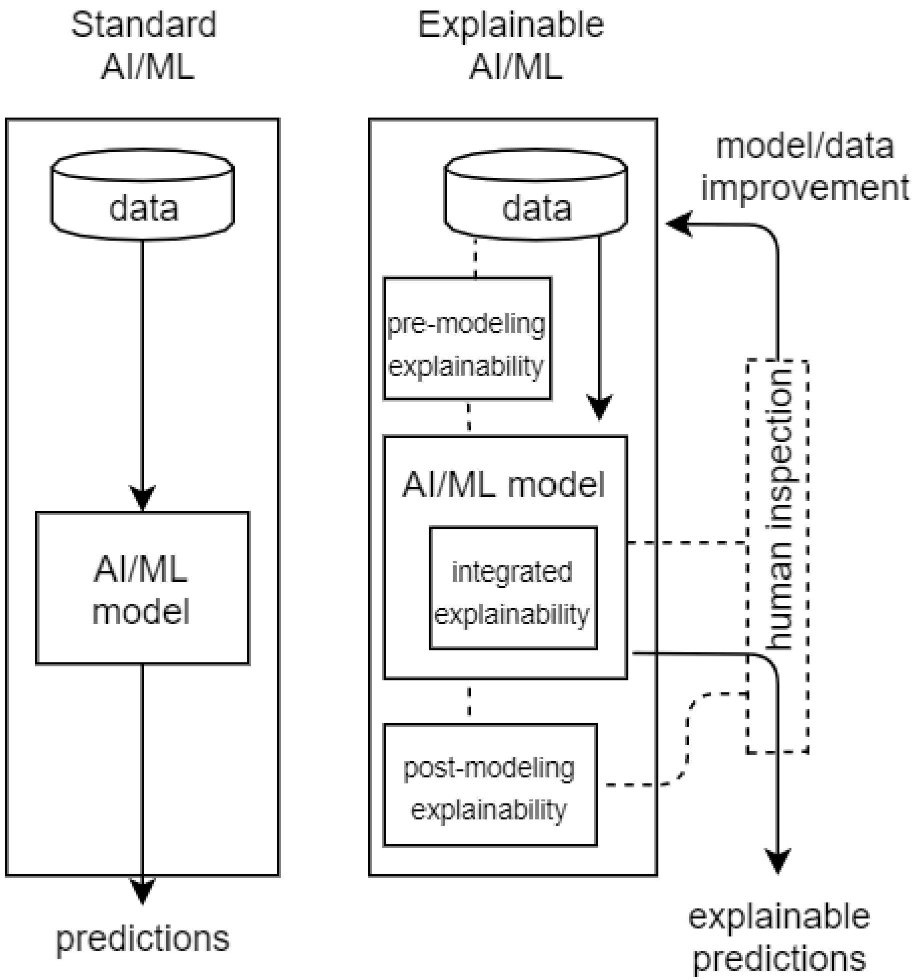
Domain Knowledge Aided Explainable and Fair Artificial Intelligence

Sheikh Rabiul Islam, Ph.D.
Assistant Professor, University of Hartford

Outline

- Explainable Artificial Intelligence (XAI)
- Motivation, and why do we need to care about XAI?
- Application of XAI in different domains
- Quantification of explainability
- Fair Artificial Intelligence
- Research accomplishments and future research plans

Explainable Artificial Intelligence (XAI)



Explainability

”Explanations are . . . the currency in which we exchanged belief”

–Lombrozo (2006)

“**Explanations** consist of **interpretations** of **how** the world works and **why**”

–Deutsch (1998)

Explainability of an AI model’s prediction is the extent of **transferable qualitative understanding** of the **relationship between model input and prediction** in a **recipient friendly** manner.

Motivation

- **IoT** and **Big Data** has necessitated the adoption of Artificial Intelligence.
- AI based “black box” models **lack explainability.**
- Lack of explainability -> **lack of trust -> ethical issues.**
- Recent laws:
  - Right of Explanation
  - Algorithmic Accountability Act
- The DARPA division of DoD is spending $2 billion on an **XAI program.**
- DoD’s Ethical Principles for AI (Feb, 2020): **responsible, equitable, traceable, reliable, and governable.**

Existing Techniques/Tools for Explainability

Method	Approx.	Inherent	Post/Ante	Agnos./Spec.	Global/Local
Partial Dependence Plot (PDP)	Yes	No	Post	Agnostic	Global
Individual Conditional Expectation (ICE)	Yes	No	Post	Agnostic	Both
Accumulated Local Effects (ALE) Plot	Yes	No	Post	Agnostic	Global
Feature Interaction	No	Yes	Both	Agnostic	Global
Feature Importance	No	Yes	Both	Agnostic	Global
Global Surrogate	Yes	No	Post	Agnostic	Global
Local Surrogate (LIME)	Yes	No	Post	Agnostic	Local
Shapley Values (SHAP)	Yes	No	Post	Agnostic	Local
Break Down	Yes	No	Post	Agnostic	Local
Counterfactual explanations	Yes	No	Post	Agnostic	Local
Adversarial examples	Yes	No	Post	Agnostic	Local
Prototypes	Yes	No	Post	Agnostic	Local
Influential instances	Yes	No	Post	Agnostic	Local

There is a lack of an explainability method, which is, at the same time **actual and direct, model agnostic, and local** method to utilize the full potential.

Research Question

- How to **make explainable and fair prediction** by keeping the core AI/ML-based prediction algorithm **unchanged**?
- How to **leverage domain knowledge** to improve explainability and fairness? Also, how that affects input and output, and what are the associated gain/compromises?
- How to **translate** the proposed approach to **a different domain**?

XAI for Finance—Bankruptcy prediction (Test Case I)

Test Case I: Infusing Domain Knowledge in AI-based "black box" Models for Better Explainability with Application in Bankruptcy Prediction

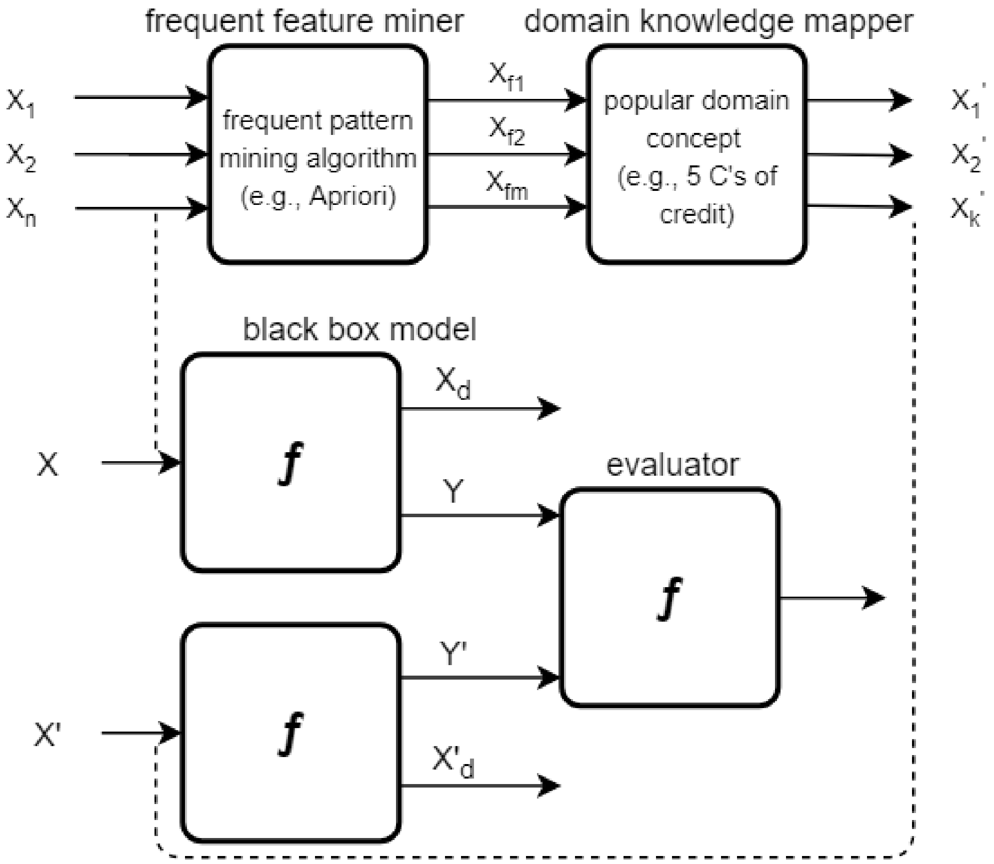


Figure 2: Feature Generalizer (on the top); Evaluator (on the bottom)

***Sheikh Rabiul Islam**, William Eberle, Sid Bundy, and Sheikh Khaled Ghafoor, "**Infusing domain knowledge in AI-based "black box" models for better explainability with application in bankruptcy prediction**", 25th* ***ACM SIGKDD**, Workshop: Anomaly Detection in Finance, 2019.*

Test Case I: Domain Knowledge Incorporation

5 C's	Mapped Feature from Frequent Feature Set
Character	creditScore, creditScoreOriginal, creditScore-Coborrower
Capacity	debtToIncomeRatioOriginal, currentDelinquen-cyStatus
Capital	UPBactual, UPBoriginal
Conditions	propertyState, interestRateCurrent, interestRa-teOriginal, postalCode
Collateral	LTV, LTVoriginal, CLTV, CLTVoriginal

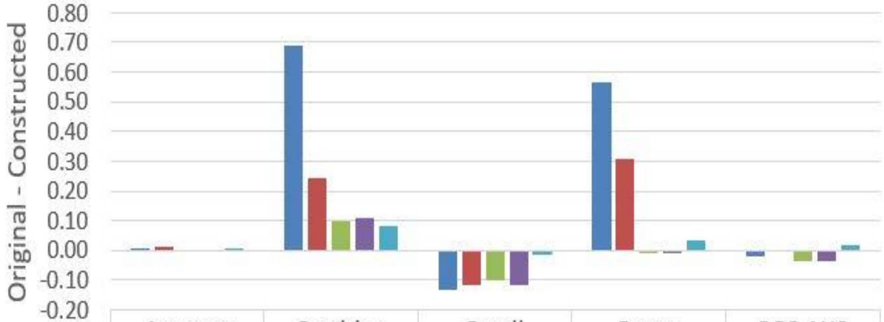
Table 2: Feature mapping with 5 C's of credit

$$C = \sum_{i=0}^{n} C_{\mathrm{i}} V_{\mathrm{i}} \quad \ldots\ldots.. \quad (1)$$

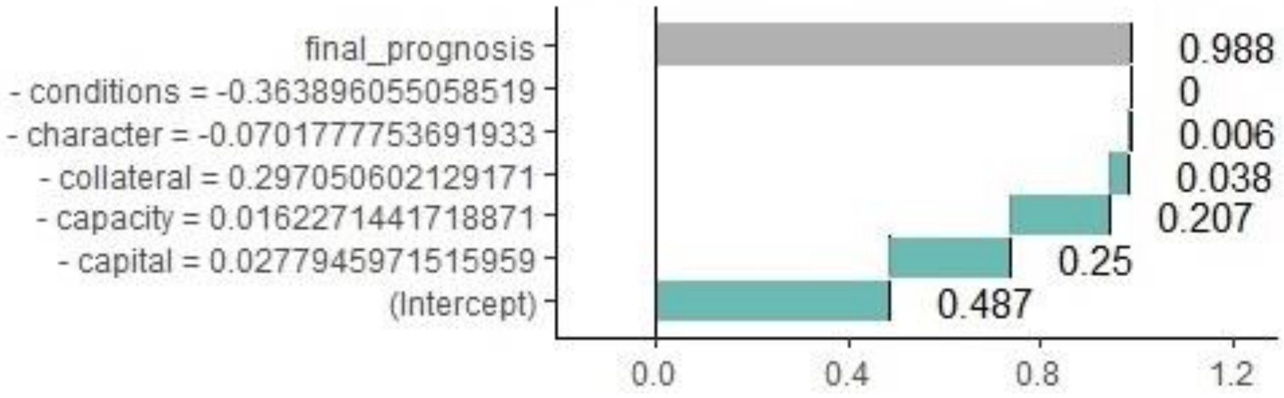
{ character, capacity, capital, conditions, collateral }

Test Case I: Results

	Accuracy	Precision	Recall	Fscore	ROC-AUC
ANN	0.00	0.05	-0.02	0.02	0.06
SVM	0.00	0.05	0.12	0.07	0.00
RF	0.00	0.02	-0.10	-0.05	-0.03
ET	0.00	0.02	0.03	0.03	-0.03
GB	0.00	-0.03	-0.02	-0.03	0.00

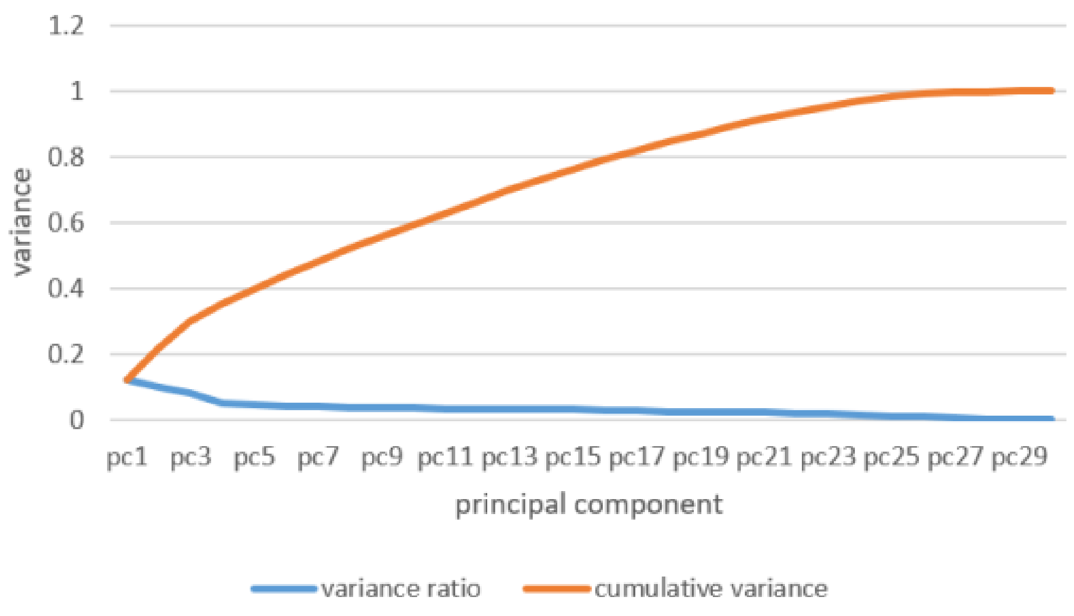


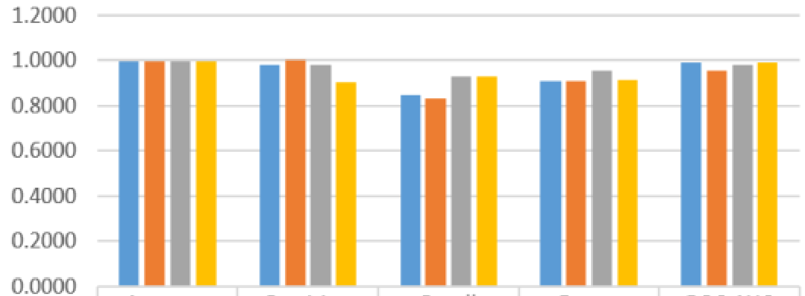
	Accuracy	Precision	Recall	Fscore	ROC-AUC
ANN	0.01	0.69	-0.14	0.57	-0.02
SVM	0.01	0.24	-0.12	0.31	0.00
RF	0.00	0.10	-0.10	-0.01	-0.03
ET	0.00	0.11	-0.12	-0.01	-0.03
GB	0.00	0.08	-0.02	0.04	0.02



Information sacrificing?

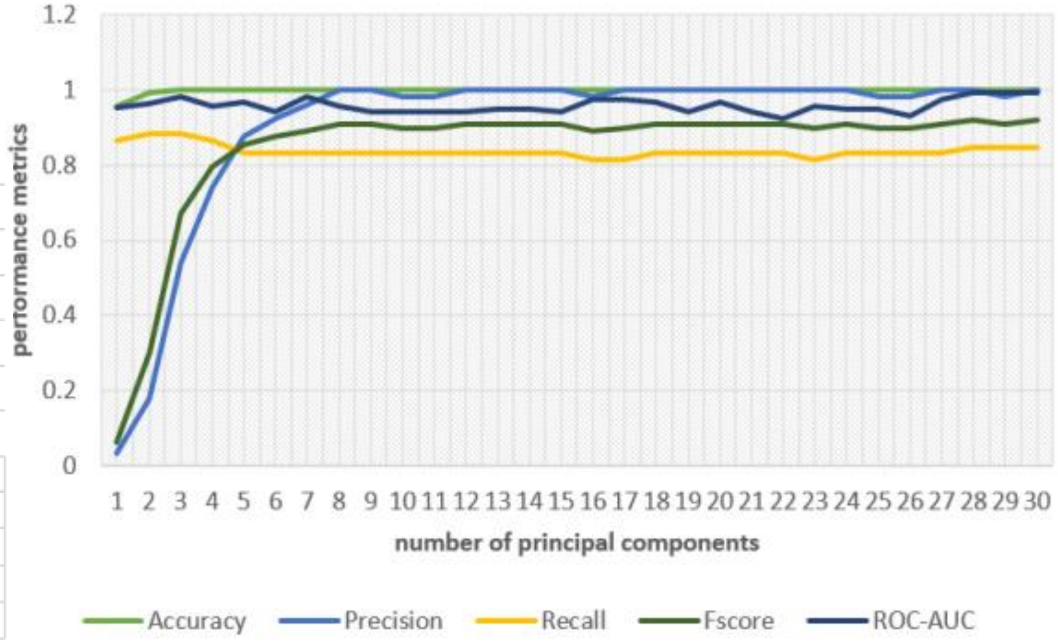
Our approach **retains enough valuable information** and does not trade valuable information to achieve explainability.





	Accuracy	Precision	Recall	Fscore	ROC-AUC
PCA	0.9997	0.9804	0.8475	0.9091	0.9914
Original	0.9997	1.0000	0.8305	0.9074	0.9576
Domain	0.9999	0.9821	0.9322	0.9565	0.9830
Constructed	0.9997	0.9016	0.9322	0.9167	0.9915

PCA Original Domain Constructed



XAI for Cybersecurity—Network Intrusion Detection and Response (Test Case II)

Test Case II: Domain Knowledge Aided Explainable Artificial Intelligence for Intrusion Detection and Response

- AI has become an **integral part** of modern-day security solutions.
- Intrusion **detection is fast**, but the **response is at human speed**.
- We infuse popular domain knowledge (e.g., CIA principle) in the model.
- **CICIDS** dataset covering popular attacks: DDoS, Brute Force, XSS, SQL Injection, Infiltration, Portscan, and Botnet.

***Sheikh Rabiul Islam**, William Eberle, Sheikh K. Ghafoor, Ambareen Siraj, and Mike Rogers,*
*"**Domain Knowledge Aided Explainable Artificial Intelligence for Intrusion Detection and Response** ",*
Accepted in AAAI-MAKE 2020.

Test Case II: Domain Knowledge Incorporation

Table 1: Mapping of network attack with related component of CIA principles

Attack	Related component of CIA
DoS GoldenEye	A
Heartbleed	C
DoS hulk	A
DoS Slowhttp	A
DoS slowloris	A
SSH-Patator	C
FTP-Patator	C
Web Attack	C, I, A
Inflitration	C
Bot	C, I, A
PortScan	C
DDoS	**A**

$$f(feature) \rightarrow attack \tag{1}$$

$$f(attack) \rightarrow C, I, or\, A \tag{2}$$

$$C = \sum_{i=0}^{n} C_{\mathrm{i}} V_{\mathrm{i}} \tag{3}$$

$$I = \sum_{i=0}^{n} I_{\mathrm{i}} V_{\mathrm{i}} \tag{4}$$

$$A = \sum_{i=0}^{n} A_{\mathrm{i}} V_{\mathrm{i}} \tag{5}$$

Table 2: Mapping of feature with related component of CIA principles

Feature	Description	In top 3 features of attack	Renamed feature
ACK Flag Count	Number of packets with ACK	SSH-Patator	ACK Flag Count - C
Active Mean	Mean time a flow was active before becoming idle	DoS Slowhttp, Infiltration	Active Mean - AC
Active Min	Minimum time a flow was active before becoming idle	DoS Slowhttp	Active Min - A
Average Packet Size	**Average size of packet**	**DDoS**	**Avg Packet Size - A**
Bwd IAT Mean	Mean time between two packets sent in the backward direction	DoS slowloris	Bwd IAT Mean - A
Bwd Packet Length Std	Standard deviation size of packet in backward direction	DoS Hulk, DoS GoldenEye, DDoS, Heartbleed, DoS Hulk	Bwd Packet Length Std - AC

Test Case II: Experimental Results

- **Doman knowledge mapped features** provides better explainability with negligible compromises (<.005%) in any performance metrics.

$$P(D) = b + \sum_{g=0}^{G} contribution(g)$$

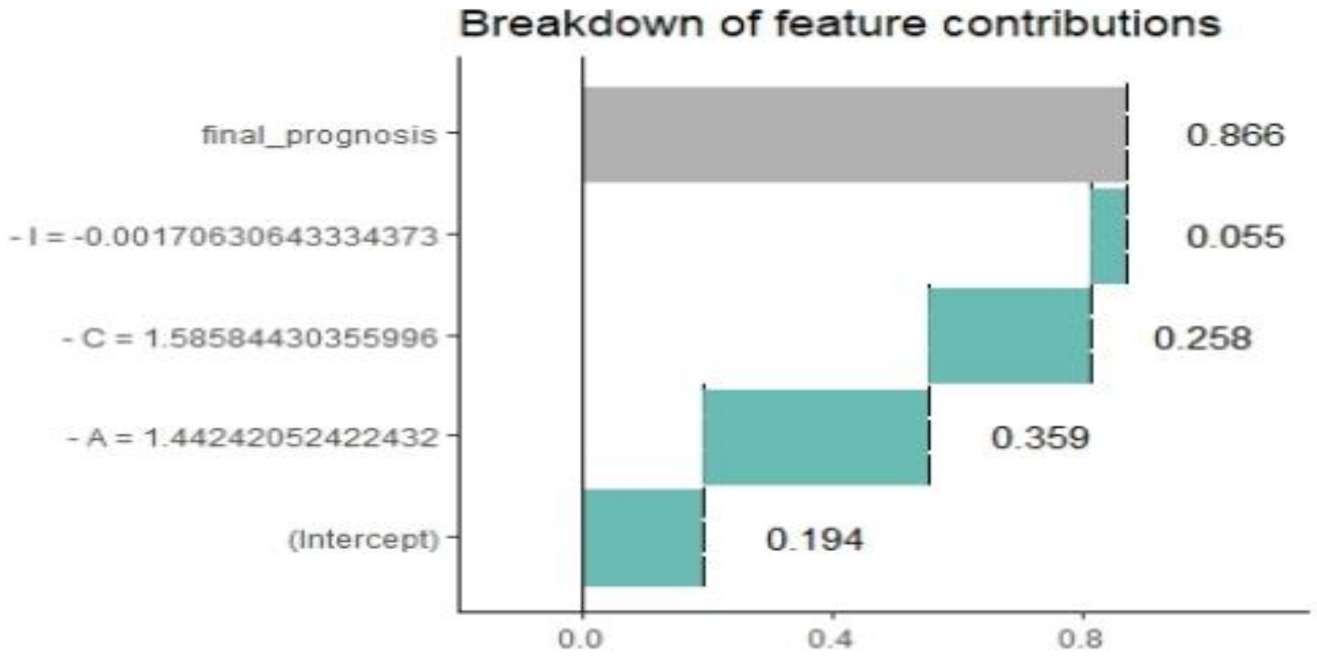
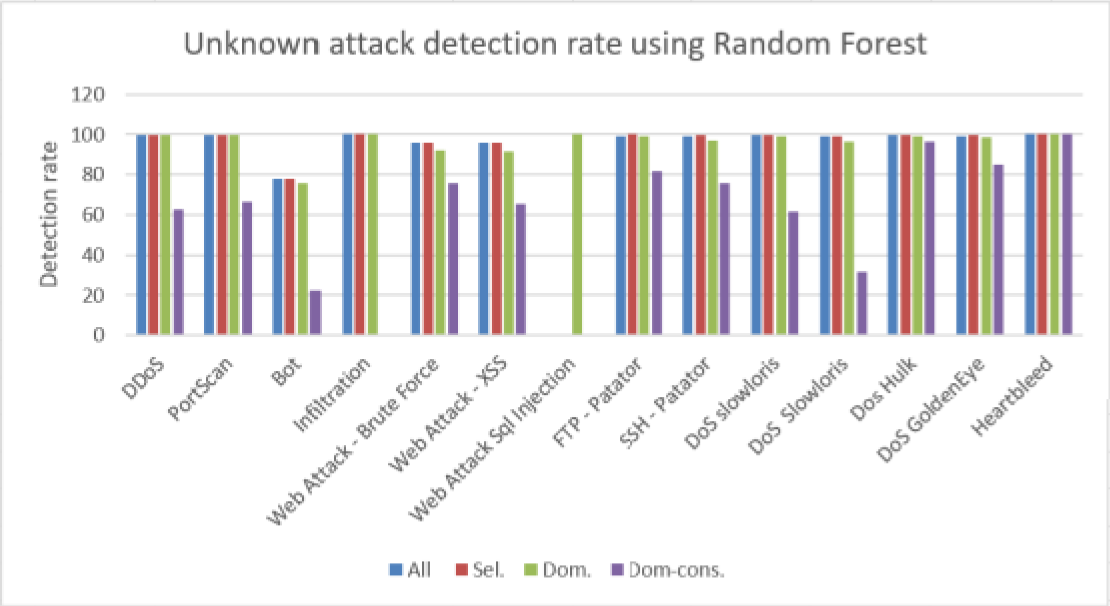
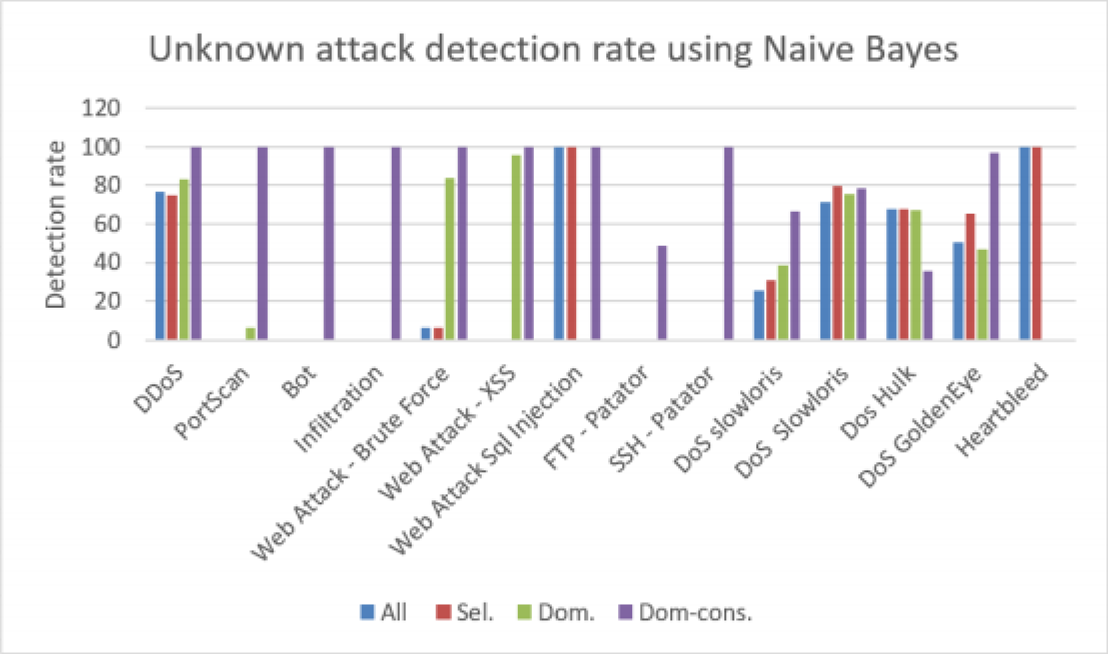


Figure 1: Breakdown of feature contributions in the prediction for a random sample.

Test Case II: Experimental Results

- The infused domain knowledge (i.e., constructed features) **generalizes the model** to work well with **unknown attack and Big Data**.





Evaluation of Explainability

- **Application-level** evaluation (real task by **domain expert**)
- **Human-level** evaluation (simplified task by a **layperson**)
- **Function-level** evaluation (**proxy task**)

Doshi-Velez, Finale and Kim, Been. Towards a rigorous science of interpretable machine learning." *arXiv preprint arXiv:1702.08608* , 2017.

Proposed Explainability Quantification Method

$$E = \frac{1}{N_c}$$

E = explainability;

N_c = # cognitive chunks;

$$E = \frac{1}{N_c} + (1 - I)$$

I = interaction;

N_i = # input cognitive chunks;

$$E = \frac{1}{N_i} + \frac{1}{N_o} + (1 - I)$$

N_o = # cognitive chunks involves in explanation representation

$$E = \frac{w_1}{N_i} + \frac{w_2}{N_o} + w_3(1 - I)$$

Sheikh Rabiul Islam, William Eberle, Sheikh K. Ghafoor. "Towards Quantification of Explainability in Explainable Artificial Intelligence Methods. " Accepted in FLAIRS-33, 2020.

Experimental Results

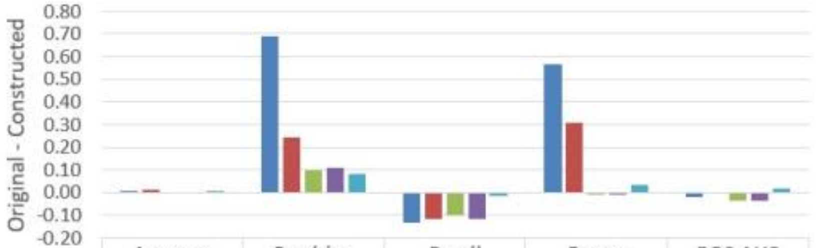
Table 1: Comparison of explainability

	Original	Domain	Constructed
Input chunks (N_i)	30	7	7
Output chunks (N_o)	30	7	5
Int. Strength (I)	0.556	0.5233	0.5251
Explainability (E)	0.1701	0.2539	0.2723

	Accuracy	Precision	Recall	Fscore	ROC-AUC
ANN	0.00	0.05	-0.02	0.02	0.06
SVM	0.00	0.05	0.12	0.07	0.00
RF	0.00	0.02	-0.10	-0.05	-0.03
ET	0.00	0.02	0.03	0.03	-0.03
GB	0.00	-0.03	-0.02	-0.03	0.00

ANN SVM RF ET GB

Figure 2: Dispersion in performance—*original features* minus *domain-related features*



	Accuracy	Precision	Recall	Fscore	ROC-AUC
ANN	0.01	0.69	-0.14	0.57	-0.02
SVM	0.01	0.24	-0.12	0.31	0.00
RF	0.00	0.10	-0.10	-0.01	-0.03
ET	0.00	0.11	-0.12	-0.01	-0.03
GB	0.00	0.08	-0.02	0.04	0.02

ANN SVM RF ET GB

Figure 3: Dispersion in performance—*original features* minus *newly constructed features*

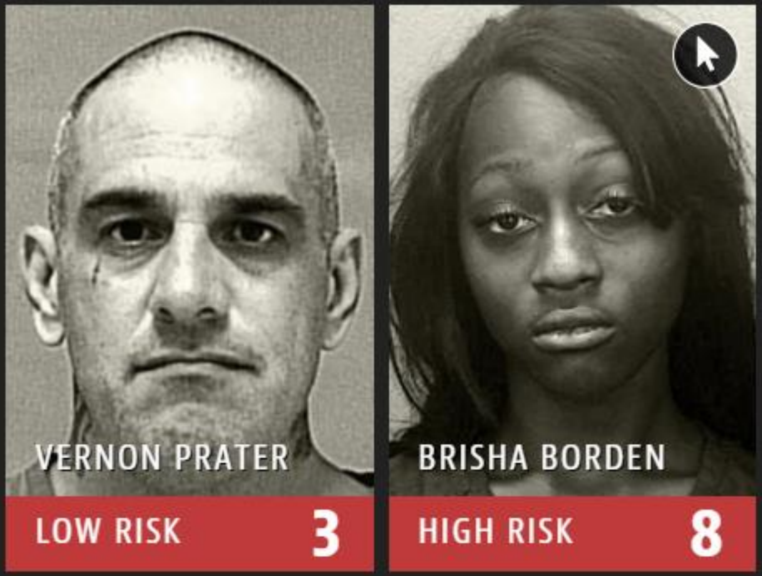
Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

Two Petty Theft Arrests



Borden was rated high risk for future crime after she and a friend took a kid's bike and scooter that were sitting outside. She did not reoffend.

Two Drug Possession Arrests

Fugett was rated low risk after being arrested with cocaine and marijuana. He was arrested three times on drug charges after that.

Two DUI Arrests

Lugo crashed his Lincoln Navigator into a Toyota Camry while drunk. He was rated as a low risk of reoffending despite the fact that it was at least his fourth DUI.

ProPublica is an independent, non-profit newsroom that produces investigative journalism in the public interest

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Predictive Policing System is Biased

- According to an NYU Law Review [1], “predictive policing systems to forecast criminal activity and allocate police resources are built on data produced during documented periods of flawed, racially-biased, and sometimes unlawful practices and policies (a.k.a. ”**dirty policing**”). As a result, decisions grounded on systematically biased data (”**dirty data**”) can not escape the legacies of unlawful and biased policing practices”

- Predictive policing has been found to lead to disproportionate patrolling and arrests in communities of color and to disproportionate ``false positives'' among people of color [2] (relative to whites) in both predicting their likelihood of future offending and in their identification as perpetrators using AI facial recognition from video surveillance evidence [3]. As a result, in June 2020, Santa Cruz, CA has become the first U.S. city to ban predictive policing [4].

Primary Factors for Fairness Related Risks

- **Data bias** and **algorithmic bias** are the primary contributing factors for fairness related risks in AI-based decision making.

- A particular group can be disproportionately represented in the data due to **natural or systematic bias** in the data collection process; and a group could be subject to statistical discrimination that is inherent in AI algorithms.

What is Bias in Decision-making?

- In the context of decision making, a fair decision is free from favoritism or prejudice towards individuals or groups based on their inherent or acquired characteristics; in contrast, a biased decision is skewed towards a particular person or group.

- Data bias and algorithmic bias are the primary contributing factors for fairness related risks.

Different Kinds of Discriminations

(1) **Direct Discrimination**: when the protected attributes (e.g., sex, race) of individuals explicitly result in a non-favorable outcome toward them;

(2) **Indirect discrimination**: when non-protected attributes (e.g., zip code) are used for decision making (e.g., loan approval decision), but the individual can still be discriminated from the implicit effect (e.g., an implicit guess of race from zip code) of the protected attribute (e.g., race);

(3) **Systemic discrimination**: results from flawed policies, custom, or behaviors (i.e., perpetuating discrimination against certain groups) that are part of the culture or structure of an organization;

(4) **Statistical discrimination**: results from the use of group statistics to judge an individual belonging to that group.

Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. arXiv preprint arXiv:1908.09635, 2019.

Different Grounds for Discriminations

(1) Title VII of the Civil Rights Act of 1964: Race/color, religion, sex, sexual harassment, pregnancy, national origin.

(2) Equal Pay Act of 1963: equal pay and compensation discrimination

(3) Age Discrimination in Employment Act of 1967.

(4) Rehabilitation Act of 1973: employment discrimination based on disability.

Can you think of any other ground for discrimination?

My students think the following could also be a ground for discrimination:

Educational background; Criminal history; Language; Cultural background; Zip code; Name; LGBTQ; Individual’s physical; Nepotism; Someone with kids.

How can we defend bias in AI?

- It is very hard to completely eliminate bias in data, and most algorithms are prone to bias.

- We have to **detect** and **mitigate** bias.

Tools to detect an mitigate bias

- Aequitas: Bias and Fairness Audit Toolkit (http://aequitas.dssg.io)
- Scikit-Fairness (https://scikit-fairness.netlify.app)
- AI Fairness 360 (https://aif360.mybluemix.net/)

Fairness in AI

Advancing Fairness in Public Funding Using Domain Knowledge

- In the transportation sector, in general, the **funding allocation** in a particular geographic area corresponds to the **population** in that area.

- However, we found that **areas with high diversity index have a higher public transit ridership**, and this is a crucial piece of information to consider for an equitable distribution of funding.

- Therefore, in our proposed approach, we use the **above fact as domain knowledge** to guide the developed model to **detect and mitigate** the **hidden bias** in funding distribution.

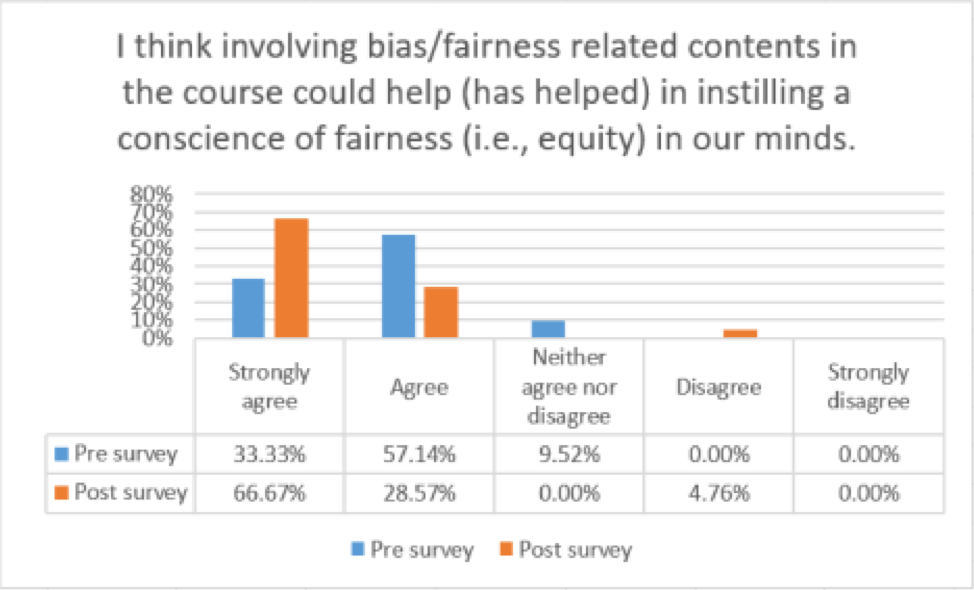
- Our intervention has the **potential to improve the declining rate of public transit ridership** which has decreased by 3% in the last decade.

Goolsby, T., Islam, S. R., & Russell, I. Advancing Fairness in Public Funding Using Domain Knowledge. AAAI Spring Symposium 2022

Fairness in AI

(1) we share our experiences with ongoing work to develop and evaluate a **cybersecurity curricular module** that demonstrates (a) data bias detection, (b) data bias mitigation, (c) algorithmic bias detection, and (d) algorithmic bias mitigation, using a network intrusion detection problem on real-world data.

(2) we share our experience in building and incorporating a fairness module [3] in **undergraduate Data Mining course.** The module includes **lectures** and **hands-on exercises**, using state-of-the-art and open-source **bias detection and mitigation** software, on real-world datasets.



	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
Pre survey	33.33%	57.14%	9.52%	0.00%	0.00%
Post survey	66.67%	28.57%	0.00%	4.76%	0.00%

(1) Islam, S. R., Russell, I., Eberle, W., & Dicheva, D. (2022, March). Incorporating the Concepts of Fairness and Bias into an Undergraduate Computer Science Course to Promote Fair Automated Decision Systems. In *Proceedings of the 53rd ACM Technical Symposium on Computer Science Education V. 2* (pp. 1075-1075).

(2) Incorporating the Concept of Bias and Fairness in Cybersecurity Curricular Module, Sheikh Rabiul Islam, Ingrid Russell, and Maanak Gupta
Accepted in SIGCSE TS 2023

(3) https://sheikhrabiul.github.io/publications/bias-fairness-module-v1.pdf

Health Informatics

[1] Visualization of COVID-19 Symptoms across different locations and times.

[2,3] Mined tweets and COVID related research paper archive to enhance the chain of infection and observing global research trends.

[4] Analyzed the impact of non-pharmaceutical Interventions (NPIs), such as Stay-at-Home, across different race and communities.

[1] Biddut Sarker Bijoy, Syeda Jannatus Saba, Souvika Sarker, Md Saiful Islam, Sheikh Rabiul Islam, Md. Ruhul Amin and Shubhra Kanti Karmaker, “COVID19-Alpha : Spatio-Temporal Visualization of COVID-19 Symptoms through Tweet Analysis”, ACM IUI '21: 26th International Conference on Intelligent User Interfaces. 2021
[2] Syeda Jannatus Saba, Biddut Sarker Bijoy, Souvika Sarkar, Md Saiful Islam, Sheikh Rabiul Islam, Md. Ruhul Amin and Shubhra Kanti Karmaker, “Towards Containing COVID-19 Pandemic by Mining Knowledge from Scientific Literature and Social Media”, 17th Int. Conference on Data Science (ICDATA’21).
[3] Souvika Sarker, Shubhra Karmaker, Mohammad Ruhul Amin, Biddut Sarker Bijoy, Yash Mahajan, Sheikh Rabiul Islam, “Ad-Hoc Monitoring of COVID-19 Global Research Trends for Well-Informed Policy Making,ACM Transactions on Intelligent Systems and Technology
[4] Yash Mahajan, Sheikh Rabiul Islam, Mohammad Ruhul Amin, and Shubhra Karmaker “Data-Driven Estimation of Effectiveness of COVID-19 Non-pharmaceutical Intervention Policies”, IEEE BigData 2022 - 5th Special Session on HealthCare Data

Summary of Research Accomplishments

- My research has resulted in ***23 research papers*** and **one book**.

- [6], [7], [9], and [15] are on domain knowledge aided **Explainable Artificial Intelligence** (XAI) research, considering the Big Data aspect (e.g., intrusion detection from the stream of IoT data);

- [9], [3], and [2] are on **predictive modeling** (e.g., bankruptcy prediction);

- [14], [8], and [5] are on **pattern discovery and anomaly/intrusion detection** (e.g., illegal insider trading detection, network intrusion detection).

- [10-13] and [20-21] are on **health informatics** focusing on **COVID-19** and Dengue.

- [17], [19], and [22] are focused on introducing a **fairness module** in the Computer Science curriculum, and [18] is focused on **advancing fairness in public funding** allocation.

- Recently, I have co-edited and authored a chapter in a **book** [23] on **Explainable Artificial Intelligence for Cyber Security.**

Future Research Agenda

- Advancing ongoing work to develop **Fair, Accountable, and Transparent** AI system for different application domain (e.g., **Cybersecurity**, **Social justice related issues**).

- XAI enhances **understandability** directly and it increases **trust** as a byproduct. **Using explanation techniques to uncover potential bias related risks** is another future research agenda.

- Human Machine Teaming
  - To ensure responsible use of AI
  - Multidisciplinary effort for the design of human-centered AI.

Research Grant

- **Internal**
  - University of Hartford: 2021-22 Greenburg Junior Faculty Research Grant
  - University of Hartford: 2021-22 Dean's Research and Teaching Grant
  - University of Hartford: 2021-22 Grants to Promote Diversity, Equity, and Inclusion within the Classroom
  - University of Hartford: 2020-21 Grants to Promote Diversity, Equity, and Inclusion within the Classroom

- **External Grant Writing Experience**
  - NSF CRII (Single PI) - Declined
  - NSF Fairness in AI (Collaborative) - Declined
  - NSF Training-based Workforce Development for Advanced Cyberinfrastructure (Collaborative) - Pending

References

[23] Mohiuddin Ahmed, **Sheikh Rabiul Islam**, Adnan Anwar, Nour Moustafa, Al-Sakib Khan Pathan

Book: *"Explainable Artificial Intelligence for Cyber Security",* Springer Nature

[22] **Sheikh Rabiul Islam**, Ingrid Russell, and Maanak Gupta, "*Incorporating the Concept of Bias and Fairness in Cybersecurity Curricular Module",* Accepted in SIGCSE TS 2023

[21] Yash Mahajan, **Sheikh Rabiul Islam**, Mohammad Ruhul Amin, and Shubhra Karmaker

"*Data-Driven Estimation of Effectiveness of COVID-19 Non-pharmaceutical Intervention Policies",*

IEEE BigData 2022 - 5th Special Session on HealthCare Data

[20] Souvika Sarker, Shubhra Karmaker, Mohammad Ruhul Amin, Biddut Sarker Bijoy, Yash Mahajan, **Sheikh Rabiul Islam**, "*Ad-Hoc Monitoring of COVID-19 Global Research Trends for Well-Informed Policy Making,*ACM Transactions on Intelligent Systems and Technology

[19] **Sheikh Rabiul Islam**, Ingrid Russell, William Eberle, Darina Dicheva, *"Instilling conscience about bias and fairness in automated decisions",* Journal of Computing Sciences in Colleges

[18] Thomas Goolsby, **Sheikh Rabiul Islam**, Ingrid Russell, "*Advancing Fairness in Public Funding Using Domain Knowledge"*,AAAI 2022 Spring Symposium (HOW FAIR IS FAIR? ACHIEVING WELLBEING AI)

[17] **Sheikh Rabiul Islam**, Ingrid Russell, William Eberle, Darina Dicheva, "*Incorporating the Concepts of Fairness and Bias into an Undergraduate Computer Science Course to Promote Fair Automated Decision Systems*", **SIGCSE TS 2022**

[16] A. N. M. Bazlur Rashid, Mohiuddin Ahmed, **Sheikh Rabiul Islam, "***A Supervised Rare Anomaly Detection Technique via Cooperative Co-Evolution-Based Feature Selection using Benchmark UNSW NB15 Dataset*", Accepted and awaiting publication in **UbiSec 2021**.

[15] **Sheikh Rabiul Islam** and William Eberle, "*Implications of Combining Domain Knowledge in Explainable Artificial Intelligence*", **AAAI-MAKE, 2021**.

[14] Hong Liu, Chen Zhong, Awny Alnusair, **Sheikh Rabiul Islam**, "*FAIXID: A Framework for Enhancing AI Explainability of Intrusion Detection Results Using Data Cleaning Techniques*", **Journal of Network and Systems Management 2021.**

References

[13] Syeda Jannatus Saba, Biddut Sarker Bijoy, Souvika Sarkar, Md Saiful Islam, **Sheikh Rabiul Islam**, Md. Ruhul Amin and Shubhra Kanti Karmaker, "*Towards Containing COVID-19 Pandemic by Mining Knowledge from Scientific Literature and Social Media*", 17th Int. Conference on Data Science (**ICDATA'21**).

[12] Nazia Tasnim, Md. Istiak Hossain Shihab, Moqsadur Rahman, **Sheikh Rabiul Islam**, Mohammad Ruhul Amin, "*Exploring the Scope and Potential of Local Newspaper-based Dengue Surveillance in Bangladesh*", Joint KDD 2021 Health Day and 2021 **KDD Workshop** on Applied Data Science for Healthcare.

[11] Biddut Sarker Bijoy, Syeda Jannatus Saba, Souvika Sarker, Md Saiful Islam, **Sheikh Rabiul Islam**, Md. Ruhul Amin and Shubhra Kanti Karmaker, "*COVID19-Alpha : Spatio-Temporal Visualization of COVID-19 Symptoms through Tweet Analysis*", **ACM IUI '21**: 26th International Conference on Intelligent User Interfaces. 2021

[10] Nazia Tasnim, Md. Istiak Hossain Shihab, Moqsadur Rahman, Jillur Rahman Saurav, **Sheikh Rabiul Islam**, "*Observing the Unobserved: A Newspaper Based Dengue Surveillance System for the Low-Income Regions of Bangladesh*", The 34th International **FLAIRS** Conference.

[9] **Sheikh Rabiul Islam**, William Eberle, Sid Bundy, and Sheikh Khaled Ghafoor. "*Infusing domain knowledge in AI-based 'black box' models for better explainability with application in bankruptcy prediction.*" *25th* **ACM SIGKDD 2019**, *Workshop: Anomaly Detection in Finance*. ACM, 2019.

[8] Qian Chen, **Sheikh Rabiul Islam**, Henry Haswell, and Robert A. Bridges. "*Automated Ransomware Behavior Analysis: Pattern Extraction and Early Detection*." The 2nd International Conference on Science of Cyber Security (**SciSec 2019**). Springer 2019.

References

[7] **Sheikh Rabiul Islam**, William Eberle, Sheikh K. Ghafoor, Ambareen Siraj, and Mike Rogers. "*Domain Knowledge Aided Explainable Artificial Intelligence for Intrusion Detection and Response.*" **AAAI-MAKE, 2020**.

[6] **Sheikh Rabiul Islam**, William Eberle, Sheikh K. Ghafoor. "*Towards Quantification of Explainability in Explainable Artificial Intelligence Methods.*" **FLAIRS-33,** 2020.

[5] **Sheikh Rabiul Islam**, Sheikh Khaled Ghafoor, and William Eberle. "*Mining Illegal Insider Trading of Stocks: A Proactive Approach.*" **2018 IEEE International Conference on Big Data (Big Data)**. IEEE, 2018.

[4] Md Mosharaf Hossain, Thomas M. Hines, Sheikh K. Ghafoor, **Sheikh Rabiul Islam**, and Ramakrishnan Kannan. "*A Flexible-blocking Based Approach for Performance Tuning of Matrix Multiplication Routines for Large Matrices with Edge Cases.*" The 2nd Workshop on Benchmarking, Performance Tuning and Optimization for Big Data Applications (BPOD), **IEEE Big Data.** IEEE, 2018.

[3] **Sheikh Rabiul Islam**, William Eberle, and Sheikh Khaled Ghafoor. "*Credit Default Mining Using Combined Machine Learning and Heuristic Approach.*" 14th Int. Conference on Data Science (**ICDATA'18**). CSREA Press, 2018.

[2] **Sheikh Rabiul Islam**, William Eberle, and Sheikh Khaled Ghafoor. "*Mining Bad Credit Card Accounts from OLAP and OLTP.*" ACM International Conference on Compute and Data Analysis (**ICCDA'17**). ACM, 2017.

[1] Tanzeem Bin Noor, Md Rounok Salehin, **Sheikh Rabiul Islam**. "*A clustering scheme for peer-to-peer file searching in mobile ad hoc networks.*" International Journal of Advanced Research in Computer and Communication Engineering. **IJARCCE**, 2012.

Thank You

Q & A

https://SheikhRabiul.github.io/

shislam@hartford.edu