

CSCE 590-1: Trusted AI

Lecture 30: Concluding Lecture

PROF. BIPLAV SRIVASTAVA, AI INSTITUTE

7TH DEC, 2021

Carolinian Creed: “I will practice personal and academic integrity.”

Organization of Lecture 30

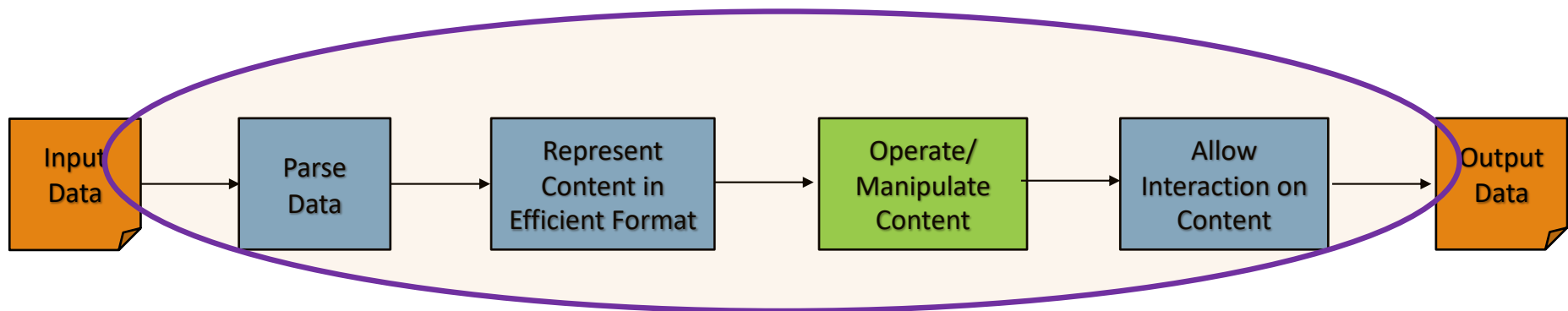
- Introduction Segment
 - Announcements
- Main Segment
 - Course Summary
 - Student Comments
- Concluding Segment
 - Ask me anything

Introductory Segment

Announcements

- Everything pending is due now
- Grades to be posted within 2 days
 - Project evaluated
 - A couple of projects are highlighted on course GitHub link

Main Segment



Learning Objectives

Undergraduate students will be able to:

L1: Explain, execute and create AI-based analytical methods to process data: (a) unstructured data, (b) semi-structured data, (c) structured data

L2: Explain AI methods in data analysis: (a) Learning methods, (b) Reasoning, (c) Representation and standardization – knowledge graphs/ ontology, (d) Preferences, (e) Handling Uncertainty

L3: Identify trust issues in AI methods: (a) fairness and bias, (b) harmful language, (c) data privacy

L4: Methods and tools to promote trust: (a) Data sampling and synthetic data, (b) Testing and rating for communication, (c) Algorithmic innovations like differential privacy and explanations

Graduate students will be able to do all of the above, and:

L5: Evaluate gaps in Trusted AI tools and create new datasets to handle them

L6: Explain emerging standards, frameworks and laws.

L7: Explain research findings in open areas and critique their contributions

What We Covered

<https://sites.google.com/site/biplavsrivastava/teaching/csce-590-trusted-ai>

Undergraduate Student Assessment

Tests	1000 points
• Course Project – report, in-class presentation	500 points
• Quiz – best of 3 from 4	300 points
• Final Exam	200 points
Total	1000 points

- Project: 40% + 10%:
 - project report (40%) and code, for elevator presentation to class (10%)
 - Data analysis project
 - Dataset must be from given catalog
 - Use analytical methods to present new insights
- Quiz: 30%
 - 4 based on preceding lectures
- Exam: 20%
- Total 100%

Graduate Student Assessment

Tests	1000 points
• Course Project – report, in-class presentation	500 points
• Quiz – best of 3 from 4	200 points
• Papers: summary, in-class presentation	200 points
• Final Exam	100 points
Total	1000 points

- Project: 40% + 10%:
project report (40%) and code, for elevator presentation to class (10%)
 - Data analysis project OR
 - Dataset must be from given catalog
 - Use analytical methods to present new insights
 - Create or explore new methods (preferred for graduate students) project
 - Problem to be discussed with instructor
 - Example: Analyze sound signals to estimate crowd
- Quiz: 20%
 - 4 based on preceding lectures
- Paper presentation: 10% + 10%
 - Research paper reading (10%) and presentation to class (10)% - Total 20%
 - Read a paper accepted at a top Data / AI conference: AAAI 2019-2021, IJCAI 2019-2021, NeurIPS 2019-2021, KDD 2019-2021, SIGMOD 2019-2021. Make a 1-page summary highlighting the key points, what you liked and what you did not. Try any code given in the paper
 - Present a 1-slide summary to class (10%)
- Exam: 10%
- Total 100%

Student Assessment

A = [900-1000]
 B+ = [870-899]
 B = [800-869]
 C+ = [770-799]
 C = [700-769]
 D+ = [670-699]
 D = [600-669]
 F = [0-599]

Tests	1000 points
• Course Project	500 points
• Quiz – best of 3 from 4	300 points
• Final Exam	200 points
Total	1000 points

Tests	1000 points
• Course Project	500 points
• Quiz – best of 3 from 4	200 points
• Papers	200 points
• Final Exam	100 points
Total	1000 points

Reference Reading

- Machine Learning Testing: Survey, Landscapes and Horizons, Jie M. Zhang*, Mark Harman, Lei Ma, Yang Liu, <https://arxiv.org/abs/1906.10742>, 2019
- Bias in Data-driven AI Systems - An Introductory Survey, Staab et al, <https://arxiv.org/abs/2001.09762>, 2020
- Fairness Definitions Explained, Sahil Verma and Julia Rubin. 2018. In *Proceedings of the International Workshop on Software Fairness (FairWare '18)*. Association for Computing Machinery, New York, NY, USA, 1–7. DOI: <https://doi.org/10.1145/3194770.3194776>

Discussion

- What Worked
 - In-class experience
 - Invited lectures
 - Projects
- What Could Be Improved
 - Have more undergraduate participation
 - Additional data types (e.g., image) and AI methods (e.g., probabilistic reasoning, deep learning)

Courses Offered Till Date

- **CSCE 590-1:** [Trusted Artificial Intelligence](#) (Fall 2021)
- **CSCE 590-1:** [From Data to Decisions with Open Data: A Practical Introduction to AI](#) (Spring 2021)
- **CSCE 771:** [Computer Processing of Natural Language](#) (Fall 2020)

Research Focus

- Regulation Intelligence
 - Using AI methods (NLP: reasoning, learning, representation) to understand regulations
 - Inform improvement in regulations to help society and AI
- Trusted AI
 - Rating AI from third party perspective
 - Transparent building of consensus among multiple stakeholders/ team members
- Advanced analytical methods – demonstrated using chatbots
 - Neuro-symbolic methods
 - Sequential decision making
 - Spatio-temporal sensor data
- Applications
 - Water, Power, Cyberspace, Teaming

Concluding Segment

Ask Me Anything
