Bachelor of Technology (Computer Science and Engineering)

Semester-VII

L-0 T-0 P-2 C-1

CSE 962 PR1: Artificial Intelligence Ethics

Course Objectives

- To Study the morality and ethics in AI
- To learn about the ethics initiatives in the field of artificial intelligence
- To Study about AI standards and regulations
- To Study about social and ethical issues of robot ethics
- To Study about AI and Ethics - Challenges and opportunities

Course Outcomes (COs): Upon completion of this unit students will be able to:

1. Understand about morality and ethics in AI
2. Understand Acquire the knowledge of real time application ethics, issues and its challenges
3. Understand the ethical harms and ethical initiatives in AI
4. Apply AI standards and regulations like AI agent, safe design of autonomous and Semi-autonomous systems
5. Understand the concepts of Robotics and morality with professional responsibilities

Articulation Matrix

(Program Articulation Matrix is formed by the strength of correlation of COs with POs and PSOs. The strength of correlation is indicated as 3 for substantial (high), 2 for moderate (medium) correlation, and 1 for slight (low) correlation)

CO/PO/PSO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2	2	1	-	-	-	-	-	-	-	-	-	-	-	-
CO2	3	2	1	-	1	-	1	-	-	-	-	-	-	-	-
CO3	1	3	2	1	1	1	1	-	-	-	-	1	2	1	1
CO4	1	2	3	2	2	1	2	1	-	-	-	1	2	1	1
CO5	3	2	1	1	1	1	1	-	-	-	-	1	2	-	1

High-3 Medium-2 Low-1

Practical List **30 Hours**

1. Recent case study of ethical initiatives in healthcare, autonomous vehicles and defense following.
   1.1 Patient Privacy and Data Security
   1.2 Autonomous Vehicles Ethical Decision-Making in Autonomous Vehicles
   1.3 Defense AI Ethics in Military Applications

2. Exploratory data analysis on a 2 variable linear regression model.

Example Dataset, Suppose we have a dataset with the following two variables:

- X: Number of hours studied 1,2,3,4,5
- Y: Test scores 50,55,65,70,75

3. Experiment the regression model without a bias and with bias solve it with help of using python.

Predicting House Prices Dataset Imagine we have a dataset with the following variables:

- Size (in square feet) 1500,1600,1700,1800,1900
- Number of Bedrooms 3,3,4,4,5
- House Price 300000,320000,340000,360000,380000

4. Classification of a dataset from UCI repository using a perception with and without bias.
   Let's classify a dataset from the UCI repository using a perceptron with and without bias. We'll use the famous Iris dataset for this example. The Iris dataset contains three classes of iris plants (Setosa, Versicolor, and Virginica) with four features: sepal length, sepal width, petal length, and petal width. For

simplicity, we'll use only two classes (Setosa and Versicolor) and two features (sepal length and sepal width).

5. Case study on ontology where ethics is at stake(developed on Ontology in AI-Powered Healthcare Systems)

6. Identification on optimization in AI affecting ethics (Bias and Fairness, Transparency and Accountability, Unethical Strategies, Privacy Concerns, Operationalizing Ethics)?

Reference(s):

1. Y. Eleanor Bird, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp and Alan Winfield ,The ethics of artificial intelligence: Issues and initiatives, EPRS | European Parliamentary Research Service Scientific Foresight Unit (STOA) PE 634.452 – March 2020
2. Patrick Lin, Keith Abney, George A Bekey, Robot Ethics: The Ethical and Social Implications of Robotics, The MIT Press- January 2014.
3. Y. Eleanor Bird, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp and Alan Winfield ,The ethics of artificial intelligence: Issues and initiatives, EPRS | European Parliamentary Research Service Scientific Foresight Unit (STOA) PE 634.452 – March 2020
4. Patrick Lin, Keith Abney, George A Bekey, Robot Ethics: The Ethical and Social Implications of Robotics, the MIT Press- January 2014.

List of e-Learning Resources:

1. https://www.javatpoint.com/robotics-and-artificial-intelligence
2. https://www.coursera.org/

Subject Tr. Academic Coordinator HoD Sr. Faculty Nominated by DOAA

Bachelor of Technology (Computer Science and Engineering)

Semester-VII

L-2 T-1 P-0 C-3

CSE 962 TR1: Artificial Intelligence Ethics

Course Objectives

- To Study the morality and ethics in AI
- To learn about the ethics initiatives in the field of artificial intelligence
- To Study about AI standards and regulations
- To Study about social and ethical issues of robot ethics
- To Study about AI and Ethics - Challenges and opportunities

Course Outcomes (COs): Upon completion of this unit students will be able to:

6. Understand about morality and ethics in AI
7. Understand Acquire the knowledge of real time application ethics, issues and its challenges
8. Understand the ethical harms and ethical initiatives in AI
9. Apply AI standards and regulations like AI agent, safe design of autonomous and Semi-autonomous systems
10. Understand the concepts of Robotics and morality with professional responsibilities

Articulation Matrix

(Program Articulation Matrix is formed by the strength of correlation of COs with POs and PSOs. The strength of correlation is indicated as 3 for substantial (high), 2 for moderate (medium) correlation, and 1 for slight (low) correlation)

CO/PO/PSO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	2	2	1	-	-	-	-	-	-	-	-	-	-	-	-
CO2	3	2	1	-	1	-	1	-	-	-	-	-	-	-	-
CO3	1	3	2	1	1	1	1	-	-	-	-	1	2	1	1
CO4	1	2	3	2	2	1	2	1	-	-	-	1	2	1	1
CO5	3	2	1	1	1	1	1	-	-	-	-	1	2	-	1

High-3 Medium-2 Low-1

UNIT I: Introduction **09 Hours**

Definition of morality and ethics in AI-Impact on society-Impact on human psychology-Impact on the legal system-Impact on the environment and the planet-Impact on trust

UNIT II: Ethical Initiatives in AI **09 Hours**

International ethical initiatives ethical harms and concern-case study: healthcare robots, autonomous vehicles, warfare and weaponization

UNIT III: AI Standards and Regulation **09 Hours**

Model Process for Addressing Ethical Concerns During System Design – Transparency of Autonomous Systems-Data Privacy Process- Algorithmic Bias Considerations Ontological Standard for Ethically Driven Robotics and Automation Systems

UNIT IV: Roboethics: Social and ethical implication of robotics **09 Hours**

Robot: robo ethics - ethics and morality -moral theories - ethics in science and technology ethical issues in an ICT society harmonization of principals - ethics and professional responsibility roboethics taxonomy

UNIT V: AI and Ethics- Challenges and Opportunities **09 Hours**

Challenges – Opportunities- ethical issues in artificial intelligence- Societal Issues Concerning the Application of Artificial Intelligence in Medicine- decision-making role in industries-National and International Strategies on AI.

45 Hours

Reference(s):

5. Y. Eleanor Bird, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp and Alan Winfield ,The ethics of artificial intelligence: Issues and initiatives, EPRS | European Parliamentary Research Service Scientific Foresight Unit (STOA) PE 634.452 – March 2020
6. Patrick Lin, Keith Abney, George A Bekey, Robot Ethics: The Ethical and Social Implications of Robotics, The MIT Press- January 2014.
7. Y. Eleanor Bird, Jasmin Fox-Skelly, Nicola Jenner, Ruth Larbey, Emma Weitkamp and Alan Winfield ,The ethics of artificial intelligence: Issues and initiatives, EPRS | European Parliamentary Research Service Scientific Foresight Unit (STOA) PE 634.452 – March 2020
8. Patrick Lin, Keith Abney, George A Bekey, Robot Ethics: The Ethical and Social Implications of Robotics, the MIT Press- January 2014.

List of e-Learning Resources:

3. https://www.javatpoint.com/robotics-and-artificial-intelligence
4. https://www.coursera.org/

Subject Tr. **Academic Coordinator** **HoD** **Sr. Faculty Nominated by DOAA**

Bachelor of Technology (Computer Science and Engineering)

Semester-VII

L-2 T-1 P-0 C-3

CSE951: Reinforcement Learning

Course Objectives

Reinforcement learning (RL) is a paradigm of learning via interactions with the environment. RL algorithms are at the frontier of current success of AI: AlphaGo, the computer program that beat humans is a RL algorithm. The objective is to provide a bottom up approach: starting from foundation in Markov decision processes (MDP), the course builds up to the state-of-the-art RL algorithms.

Course Outcomes (COs)

1. Learn how to define RL tasks and the core principles behind the RL, including policies, value functions, deriving Bellman equations
2. Implement in code common algorithms following code standards and libraries used in RL
3. Understand and work with tabular methods to solve classical control problems
4. Understand and work with approximate solutions (deep Q network based algorithms)
5. Learn the policy gradient methods from vanilla to more complex cases.

Articulation Matrix

(Program Articulation Matrix is formed by the strength of correlation of COs with POs and PSOs. The strength of correlation is indicated as 3 for substantial (high), 2 for moderate (medium) correlation, and 1 for slight (low) correlation)

CO/PO/PSO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3
CO1	3	3	3	2	-	-	-	-	-	-	-	-	1	-	-
CO2	1	2	2	1	1	-	-	-	-	-	-	-	2	1	-
CO3	2	3	3	2	3	-	-	-	-	-	-	-	1	2	-
CO4	3	2	2	1	2	-	-	-	-	-	-	-	2	1	-
CO5	3	3	3	2	3	-	-	-	-	-	-	-	1	1	-

High-3 Medium-2 Low-1

UNIT I: Introduction and Basics of RL — 09 Hours

Introduction and Basics of RL, Defining RL Framework and Markov Decision Process, Polices, Value Functions and Bellman Equations, Exploration vs. Exploitation, Code Standards and Libraries used in RL (Python/Keras/Tensorflow).

UNIT II: Tabular methods and Q-networks — 09 Hours

Tabular methods and Q-networks, Planning through the use of Dynamic Programming and Monte Carlo, Temporal-Difference learning methods (TD(0), SARSA, Q-Learning) . Deep Q-networks (DQN, DDQN, Dueling DQN, Prioritized Experience Replay).

UNIT III: Policy optimization — 09 Hours

Introduction to policy-based methods, Vanilla Policy Gradient, REINFORCE algorithm and stochastic policy search, Actor-critic methods (A2C, A3C), Advanced policy gradient (PPO, TRPO, DDPG).

UNIT IV: Model based RL — 09 Hours

Model-based RL approach

UNIT V: Recent Advances and Applications — 09 Hours

Meta-learning, Multi-Agent Reinforcement Learning, Partially Observable Markov Decision Process, Ethics in RL, Applying RL for real-world problems.

Text Books:

1. Richard S. Sutton and Andrew G. Barto, "Reinforcement learning: An introduction", Second Edition, MIT Press, 2019
2. Li, Yuxi. "Deep reinforcement learning." arXiv preprint arXiv:1810.06339 (2018).
3. Wiering, Marco, and Martijn Van Otterlo. "Reinforcement learning." Adaptation, learning, and optimization 12 (2012).

Reference Books:

1. Russell, Stuart J., and Peter Norvig. "Artificial intelligence: a modern approach."Pearson Education Limited, 2016.
2. Goodfellow, Ian, Yoshua Bengio, and Aaron Courville. "Deep learning." MIT press, 2016.
3. David Silver’s course on Reinforcement Learning (link)