

Final Review Topics, CSCI 450 – AI

Tentative point breakdown: 10-15% intro material, 15-20% general machine learning, 15-20% decision trees, 15-20% Probabilistic learning/Bayesian networks, 10-15% neural networks, 10-15% SVMs, 5-10% NLP, 5-10% Robotics

Types of questions: short answer, definitions, short essay, problem solving (planning, Bayesian networks, decision trees, sample CFGs)

1. Introductory lecture(s) (I'm not kidding – this is the one cumulative piece)
 - a. Definitions of AI (matrix)
2. Planning
 - a. Planning as search
 - b. Plan operators
 - c. State progression and goal regression
 - d. Partial order planning
3. Uncertainty and probabilistic reasoning
 - a. Representing uncertain knowledge with probabilities
 - b. Prior probability and conditional probability
 - c. Bayes' Law
 - d. Diagnostic versus causal probabilities
 - e. Conditional independence and Bayesian networks – combining evidence, understanding what nodes influence others in the network and under what conditions
 - f. Naïve Bayesian networks vs full Bayesian belief networks
4. Machine Learning
 - a. General
 - i. Supervised vs unsupervised
 - ii. Overfitting
 - iii. Variance and bias
 - iv. Training, testing and validation sets
 - v. Error rates and sources of errors
 - vi. Regression vs classification
 - b. Decision trees
 - i. Strengths and weaknesses
 - ii. Choosing split attributes (entropy and info gain)
 - iii. Algorithm for building trees
 - iv. Problems (complications) and ways of handling them
 - v. Assessing learning performance
 - vi. Nonlinear models
 - c. Random Forests/Ensemble Methods
 - i. Why ensemble methods are effective
 - ii. How trees in random forests are built and used (bagging, limiting features for splitting decisions, majority vote for decisions, etc.)
 - d. KNN (nearest neighbor)
 - i. Overall concept
 - ii. Differences from other techniques (no model learned)
 - e. Network learning (neural nets)
 - i. General concepts and definitions
 - ii. Artificial vs. biological neurons
 - iii. Activation functions, multilayer networks

- iv. How “learning” takes place
 - v. Back propagation
 - f. Support Vector Machines
 - i. Maximal margin classifier (definition, when is it applicable...)
 - ii. Support vector classifier (definition, when applicable, tuning parameter C...)
 - iii. Support vector machine (definition, when applicable)
 - iv. Kernel “trick” – what is it, why do we use it
 - v. Pros and cons of SVMs
- 5. Natural Language Processing
 - a. Syntax vs. semantics vs pragmatics
 - b. Generation vs understanding, phases of processing
 - c. Context free grammars
 - d. Challenges – why aren’t we there yet?
- 6. Robotics
 - a. The 3 D’s
 - b. Paradigms – Hierarchical vs. Reactive
 - c. Biological foundations for the reactive paradigm – behaviors, reflexes, etc.