**CSCI 446 – ARTIFICIAL INTELLIGENCE**

**EXAM 2 STUDY OUTLINE**

**Probability**

I. Uncertainty

II. Random Variables

III. Probability Distributions

A. Joint and Marginal Distributions

B. Conditional Distributions

1. Normalization Trick

V. Product Rule

VI. Chain Rule

VII. Bayes Rule

VIII. Inference

1. Enumeration

2. Bayes Rule

IX. Independence

**Markov Models**

I. Markov Models

A. Reasoning over Time or Space

B. Implied Conditional Independencies

C. Forward Algorithm

D. Stationary Distributions

II. Hidden Markov Models

A. Inference

1. Passage of Time

2. Observation

III. Particle Filters

A. Sampling

B. Elapse Time

C. Observe

D. Resample

IV. Dynamic Bayes Nets

V. Most Likely Explanation

A. Forward Algorithm vs. Viterbi Algorithm

**Bayes Nets**

I. Representation

A. Graphical Representation

1. Topology

2. Local Conditional Probability Tables

B. Semantics

1. Causality vs. Influence

C. Probabilities in Bayes Nets

II. Independence

A. Independence Assumptions

B. D-Separation

1. Causal Chains

2. Common Cause

3. Common Effect

III. Inference

A. Enumeration

B. Variable Elimination

1. Factors

2. Ordering of Variables

IV. Sampling (Approximate Inference)

A. Prior Sampling

B. Rejection Sampling

C. Likelihood Weighting

D. Gibbs Sampling

**Decision Diagrams**

I. Decision Networks

A. Calculating Maximum Expected Utility (MEU)

II. Value of Information

A. Calculating VPI

**Naïve Bayes**

I. Classification

II. Model Based Classification

A. Inference in Naïve Bayes

III. Training and Testing

A. Datasets

1. Training

2. Hold Out

3. Testing

IV. Generalization and Overfitting

V. Smoothing

A. Maximum Likelihood

B. Unseen Events

C. Laplace Smoothing

D. Linear Interpolation

VI. Tuning

VII. Features

A. Using Errors to Find Additional Features

**Perceptrons**

I. Error Driven Classification

II. Linear Classifiers

A. Feature Vectors

B. Weight Vectors

III. Weight Updates

IV. Improving the Perceptron

A. MIRA

B. Support Vector Machines

**Kernels and Clustering**

I. Case Based Learning

A. Similarity Functions

1. Invariant Metrics

B. Parametric vs. Non-Parametric Models

C. Nearest Neighbor

II. Kernelization

A. Dual Perceptron / Kernelized Perceptron

B. Kernel Trick

III. Non-Linearity

A. Non-Linear Separators

IV. Clustering – Unsupervised Learning

A. k-Means

B. Agglomerative