CS 344 – Guide 4

- Uncertainty
  - Conditional probability:
    - The measure of the probability of an event given that another event has occurred.
    - P(A|B) = the conditional probability of A given B or the probability of A under the condition B.
  - Unconditional probability:
    - The independent chance that a single outcome results from a sample of possible outcomes.
    - The likelihood that an event will take place independent of whether any other events take place or any other conditions are present.
  - Joint probability distribution:
    - Definition: Given random variables that are defined on a probability space, the joint probability distribution for those variables is the probability distribution that gives the probability that each of the random variables falls in any particular range or discrete set of values specified for that variable.
    - Probability density functions (PDF):
      - Also known as the density of a continuous random variable.
      - A function whose value at any given sample (or point) in the sample space (The set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample.
    - Joint probability density functions:
      - For continuous random variables, it is a probability density function associated to the set as a whole.
      - Defined as a function of the n variables, such that, for any domain D in the n-dimensional space of the values of the variables, the probability that a realization of the set variables falls inside the domain D is:
      - $\Pr(X_1, \ldots, X_n \subset D) = \int_D f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\, dx_1 \cdots dx_n.$
    - Need to add a new dimension to the table every time you add a new variable.
  - Bayes' theorem:
    - Describes the probability of an event, based on prior knowledge of condition that might be related to the event.
    - Bayesian inference:
      - A method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available.
      - P(H|E)=P(E|H)*P(H)/P€
  - A Plan for Spam:
    - What is the typical first approach to SPAM filtering?
      - To try to write software that recognizes individual properties of spam.
      - A.k.a. looking for specific keywords, phrases, or sentence structures.
    - What language does Graham use to specify his filtering algorithm?
      - Traditional LISP – a functional programming language.
    - What is the advantage of the statistical approach?

- One great advantage is that you don't have to read so many spams – don't have to read to recognize individual spam features and try to get into the mind of the spammer.
- Real advantage of Bayesian approach is that you know what you're measuring – assigns an actual probability rather than a "score" to e-mail, which is what feature-recognizing filters do.
- Bayesian approach considers all the evidence in the e-mail, good and bad.

❖ Pre-requisites and Pre-work:

➢ ^_^