Weekly report of lessons

Name: Suryam Arnav Kalra
Roll No: 19CS30050
The week: 23rd August 2021 to 27th August 2021

The topics covered:

- Candidate-Elimination Algorithm: Convergence, Next TE to choose, Classification of new cases using VS
- Effect of Incomplete Hypothesis Space
- Inductive Bias : Unbiased Learners, Equivalent Deductive System and IB of specific Algorithms
- Computational Complexity of VS
- PAC Learning Model
- Sample Complexity of a Learner
- Theorem of ∈-exhausting the VS
- Sample Complexity of infinite H
- Handling Noise in Data, Effect of Inductive Bias, Matching Complexities, Occam's Razor, Triple Trade-Off, Model Selection
- Decision Tree: Top-Down Construction and Principle of Construction
- Entropy
- Information Gain

Summary topic wise:

- <u>Convergence of Candidate-elimination algorithm</u>: Convergence is guaranteed if there are no errors in the TE and there is an h in H describing the target concept c.
- <u>Next TE to choose</u>: Such a TE should be chosen which reduces maximally the number of hypothesis present in the VS.
- <u>Classification of new cases using VS</u>: A voting procedure can be used to classify new cases where the confidence is in the majority vote.
- <u>Effect of Incomplete Hypothesis Space</u>: Due to incomplete hypothesis space all the target concepts cannot be represented by some h in H. All the preceding algorithms done till now assume that c can be represented by some h in H.
- <u>Unbiased Learner</u>: These learners have no limits on the representation of hypothesis.
- <u>Inductive Bias</u>: It is the minimal set of assertions B used to logically infer the value c(x) of any instance x for any target concept c and training examples D.
- <u>Inductive Bias as an equivalent Deductive System</u>: Inductive bias is made explicit in an equivalent deductive system that logically produces the same output.
- <u>Inductive Bias of Specific Algorithms</u>: Rote Learners => NO IB , CE Algorithm => The target concept c can be represented in H, Find-S => The target concept c can be represented in H and all instances that are not positive are negative.
- <u>Computational complexity of VS</u>: The S set is linear in the number of features and TE, while the G set is exponential in the number of TE.
- <u>PAC Learning Model</u>: A consistent hypothesis having a maximum true error ∈ with respect to a population distribution D. We require only that the learner probably learns a hypothesis that is approximately correct.

- <u>Sample complexity of a learner</u>: The number of training examples required to get a successful hypothesis with a high probability.
- <u>Theorem of ∈-exhausting the VS</u>: P(VS is not ∈-exhausted) <= $|H|e^{-\in m}$
- <u>Sample complexity of infinite H</u>: It is represented by the Vapnik-Chervonenkis (VC) dimension which is the size of the largest finite subset in X shatter by H.
- <u>Handling noise in data</u>: The noise in data can come from imprecision of measurement, error in labeling or due to hidden attributes and can be handled using data cleaning processes.
- <u>Effect of inductive bias</u>: Low training error still may provide high errors on unseen inputs and the higher the number of training examples better is the model fitting.
- <u>Matching Complexities</u>: Low Complexity => Higher training and generalization error, may result in underfitting, High Complexity => Low training error, but may have high generalization error, may result in overfitting.
- <u>Occam's Razor</u>: Given comparable empirical error, a simple model would generalize better than a complex model.
- <u>Triple trade-off</u>: The trade-off between the three factors: complexity of hypothesis, amount of training data, generalization error in any data driven algorithm is the triple trade-off.
- <u>Model Selection</u>: Split the input data into 3 sets: training, validation and test sets and increase the complexity of the model keeping the training and validation errors low.
- <u>Decision Tree</u>: They can be used to represent concepts which can be represented as disjunction of conjunction of attribute literals.
- <u>Top-Down Construction</u>: Start with an empty tree, and while the training examples are not perfectly classified split the best decision attribute for the next node.
- <u>Principle of Construction</u>: Treat the entire data set as a single box and split the box with such an attribute which reduces its impurity by maximum amount.
- <u>Entropy</u>: $E = \sum -p \log(p)$
- <u>Information Gain</u>: $Gain = E(S) - \sum \frac{|S_v|}{|S|} E(S_v)$

Concepts challenging to comprehend:

PAC Learning Model along with the theorem of ∈-exhausting the VS are a little bit challenging to comprehend.

Interesting and exciting concepts:

Decision Tree, Entropy and Inductive bias are quite interesting and exciting to learn.

Concepts not understood:

After going through the book and the video lectures the concepts are clearly understood.

Any novel idea of yours out of the lessons:

Decision Trees can be used to analyze fully the possible consequences of a decision by providing us a way to quantify the values of outcomes and the probabilities of achieving them. They can be used to solve complex problems easily as they break down or branch the problem into simpler problems. They can also be used in biology for genetic analysis.