AI Foundations
Winter Holiday 2022
LinkedIn

Thinking Machines 28/12/22

- **Symbolic systems approach AI** is an approach that trains Artificial Intelligence (AI) the same way human brain learns. It learns to understand the world by forming internal symbolic representations of its "world". Symbols play a vital role in the human thought and reasoning process.
    - Like the general problem solving machine
- **Strong vs weak AI:** strong/general AI acts like a human being i.e. does everything vs weak/narrow is specific task orientated
- **Expert systems:** AI software that uses knowledge stored in a knowledge base to solve problems that would usually require a human expert thus preserving a human expert's knowledge in its knowledge base. They can advise users as well as provide explanations to them about how they reached a particular conclusion or advice. From <https://www.geeksforgeeks.org/expert-systems/>
    - Creates long list of items to make matching patterns
    - Problems w/ this: as all symbolic systems lead to combinatorial explosions
        - Too many symbols, too many matching patterns
- **Planning AI:** uses heuristic reasoning
    - An informal and speculative procedure that leads to a solution in some cases but not in others
    - Thus rules out some pattern matching; also called limiting the search space
- **Artificial Neural Network:** a system that is designed to mimic the structure of the human brain
    - Neurons in AI are organized into layers. Go from input to output layers and in between have hidden layers
    - Deep learning is when neural network had many hidden layers

    - Clustering: More sophisticated deep learning networks cluster neurons. This allows the network to create categories and effectively sort new information.
    - Key benefit of AI neural networks: Neural networks can train themselves to understand inputs and recognize those inputs when addressing big data sets.
- FINDING THE RIGHT APPROACH
    - ML > Symbolic reasoning bc ML can process extremely large bulks of data
        - Symbolic reasoning = abstract problem, but know steps
        - ML = look for patterns
        - Can mix the both, use symbolic to create constrains and ML to experiment w patterns
        - Symbolic reasoning = long setup and no outside data
        - ML = lots of data that need tweaking and updating
    - Supervised learning: training set, train and the categorize
    - Unsupervised learning: feed data and then ask AI to categorize into arbitrary categories
- **Backpropogation(of errors)/(backprop)**: use gradient ascent/descent
    - Use to adjust weights of neural networks to rectify errors
- Regression analysis
    - Regression vs classification
- **Natural Language Processing**
    - Want to communicate with machine in the normal human way
- ML vs **Data Mining**
    - Big diff is in the technology used to find the insights using data
    - ML: requires training and then using a ML framework
        - In ML, train machines the find patterns
    - Data Mining: uses broader tools w/o required training
        - Just dig thru data to find insights
- Use data science to understand the reasoning behind the answers and use neural networks to just find patterns w/o any reasoning

Machine Learning 28/12/22

- data - test data & training data

- Binary classification: classifies data into two types
- ML uses AI to find patterns in data
- Types of ML: supervised, semi-supervised, unsupervised
    - Training is the process by which the rule is developed.
    - Sorting and editing data is a preliminary activity to the whole learning process.
    - After implementing a rule, observe feedback
- Supervised Learning: tagged “labeled sample data”i.e. Training set and outputs
    - Can use ML algos or statistical regression
    - Machine takes learning from training set and applies it to test data to see if it works there too
- Unsupervised Learning:
    - Multiclass Classification: data sorted into several different groups
- Semi-supervised Learning:
    - Inductive Reasoning: using specific cases to determine general outcomes
    - Transductive Reasoning: uses context to make better conclusions abt the data
    - Transduction is reasoning from observed, specific (training) cases to specific (test) cases. In contrast, induction is reasoning from observed training cases to general rules, which are then applied to the test cases
    - Therefore, transduction uses more information and produces more specific rules than induction
    - Disadvantage of semi-supervised: can lead to large errors or confusion as transduction and induction can both be v misleading
- **Reinforcement Learning:** machine iterates to get a more accurate outcome
    - **Q-Learning:**
        - States = set environments
        - Actions = responses
        - Quality = Q
            - Determining Q:
                - Q=0 (initial)
                - Q++ (based on further improvement each time)
- Identifying fundamental rules is a product of any type of learning

- Key feature that distinguishes supervised ML from other types is that it uses labled data for training i.e. give a lot of prior knowledge abt the training data
    - A partial difference is that it classifieds variables as dependent or independent

PROBLEMS THAT USE ML

- Categories of Supervised Learning
    - Binary (all binary is supervised learning)
    - Multiclass (categorical)
    - Regression (continuous)
- **Decision Trees**
    - For binary classification challenges
    - Set up *predictors* ("questions" organised hierarchically as a tree) and the outcome
    - Root node and *decision nodes* (children nodes)
    - There must be a clear path to the yes or no outcomes i.e. low entropy
    - If a tree has too much entropy, add or substitute predictors as one or the more of the predictors is not acting efficiently
- **k-nearest neighbor (k-NN)/ lazy learning**
    - For multiclass, supervised
    - An instance-based ML algo
        - Uses a lot of computation for every instance i.e. is a brute force matching method
    - Run all computations in one big instance
    - Dis: uses a large amount of computational power - difficult to use on v large data sets
    - Aim the reduce the Euclidean distance btw the 2 data points
    - Use predictions to put as labels on the graph and look at the Euclidean distance
- **K-mean clustering**
    - Unsupervised ML algo
    - Also lazy learning, instance based
    - K stands for number of clusters

    - *Centroid* and algo looks for data points with the shortest distance with the centroid
    - Centroid is initially chosen randomly
    - Machine will iterate until it gets the optimal centroid
    - Problem w overlap of data; “high overlap of data”
    - Problem 2: sensitive to outliers; will cluster them anyways
- Dis of both K algos, bc running in one big splash, if data changes, need to rerun whole program
- In cluster analysis, an outlier is not close to any centroid and the outlier is forced unto the nearest cluster even tho its not a good fit
- Regression Analysis:
    - Look at relationship bw predictors/regressors/input variables/ independent variables and ur outcome
    - Supervised
    - Linear Reagression
        - trendline/ hyperplane
    - Regression methods that are based on statistical predictions such as linear regression are not considered good examples of ML
- **Naive Bayes**
    - Looking to see how one things influences other
    - Based on the Bayes’ theory of statistics
    - Bayesian Algo
        - Naive bayes is one of the Bayesian algos
        - Naive bc assumes all predictors are independent of one another (even thought predictors may not be independent of one another e.g. height and weight)
    - For binary, multiclass
    - Have classes and predictors
    - Class predictors probability: looks at each independent predictor and creates a probability for each class

		Terrier	Hound	Sport
Hair		0.4	0.1	0.5
Height		0.2	0.1	0.7
Weight		0.1	0.05	0.85

    - Also use weighted multiplier to decide which predictor is most predictive

APPLYING ALGOS

- Challenges: Ways of measuring the difference btw prediction and outcome
    - Bias: gap bw predicted value and the actual outcome
    - Variance: when predicted values are scattered all over the place
- Fitting the data:
    - Underfitting: too simple so doesn't fit the data
    - Overfitting; too complex even tho it fits all the data
    - Noise: natural variation in data that might not offer any insights
    - Signal: smt used to make predictions
- Selecting the best algo:
    - Labled data - supervised
    - Unlabled data - unsupervised
    - Lots of unlabeled data - k-means clustering
    - Lots of labled data - k-NN/ decision trees
    - Ensemble Modeling: (a group of items viewed as a whole rather than individually.) create ensembles of diff ML algos
        - Bagging: create several versions of the ML algo; look for best or average for inconsistent results
        - Boosting: use several diff algos to boot accuracy; like combining diff algos to together by passing the results of one into the other
        - Stacking: use several diff ML algos and stack them to improve accuracy

ML CHALLENGES

- Should never mix training and test data; always let the machine work on new untouched data
- Ask questions to accurately understand wants before starting to build anything
- Don’t make presentations w training data to prevent misunderstandings on the model’s accuracy