# SUMMARY OF DLGN DECISION TREE PAPER

R NAVEEN PRAKASH

CS23M052

This paper proposed a deep neural network model for getting hyperplanes for the Oblique Decision Tree Classification problem. The work can also be viewed as study of feature learning in deep learning. An oblique decision tree is a decision tree in which the conditions used to split the data put a constraint over a linear combination of the features. Then paper provides an example of how the greedy decision making not be able to give the correct root node hyperplane and says only techniques which can do feature learning can perform well on these types of data. Paper then gives a new metric to evaluate the ODT splits based on how discontinuous the label function around the hyperplanes denoted by ϒ(w,b) which can take value between 0 and 1. This equals to 0 means there is max continuity and equals to 1 means max discontinuity. And it is also noted that HDS increases as the distance between the node and root node decreases. Therefore, the highest HDS goes to the root node. Then the paper gives the architecture of the DLGN algorithm. The output of the DLGN is described by $\hat{y} = \sum_{\pi} f_{\pi}(x) g_{\pi}$ , where both $f_{\pi}$ and $g_{\pi}$ are learned using deep learning architecture using the notion of paths. Then the paper gives some functions to plot the decision trees given by DLGN and also compares the performance of DLGN with other algorithms. It performs better than SVM, CART. And it performs similarly to soft Decision tree algos on small dimension data and performs better with higher dimensional data. ReLU performs marginally better than DLGN. But ReLU has a lot of hyperparameter tuning which is a hassle. DLGN also gives hyperplanes that very close to true ODTs thereby performing better when the greedy method fails.

## GRADES

| | |
|---|---|
| LARP | A |
| PRML | S |
| DL and RL | Ongoing |