**Clickstream project**

ID3 decision tree learner is implemented in Python.

First, decision tree is constructed based on the training set. Chi-squared stopping criterion is used for stopping the tree construction based on p-value threshold given.

**The ID3 algorithm:**

Recursive ID3 algorithm is implemented. Attribute selection method is based on maximum gain among the possible attributes at that level. This is calculated by subtracting the entropy of the current attribute with the sum of entropies of all the other candidate attributes.

For each outcome of the splitting criterion ID3 is called recursively. We are splitting the data into two sets based on the mid value of the maximum attribute value.

**Returning condition:**

The returning condition in the algorithm is when the data tuples are pure i.e, all the data tuples are of the same class or if the attribute list is empty.

**Accuracy:**

The accuracy with different p-value thresholds is illustrated in the following graph.

**Time taken:**

It is clear from the following graph that the time taken to run the model increases with increase in p-value threshold as it involves creating larger tree.

**Memory analysis:**

Number of nodes increases with the increase in p-value threshold as lesser the p-value threshold the quicker the stopping of tree construction.