**AD OR NON-AD PREDICTION**

**NIKHIL SARAOGI**

Master of Computer Application [MCA]

**VIT UNIVERSITY CHENNAI, INDIA**

**(nikhilsarawgi9616@gmail.com)**

Under guidance of

**Dr. SATHIS KUMAR B.**

School of Computer Science and Engineering

**VIT UNIVERSITY CHENNAI, INDIA**

**(sathiskumar.b@vit.ac.in)**

**Abstract**

The dataset information comes from the paper [Learning to remove Internet advertisements](http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.5686). Number of possible advertisements on Internet pages is showed by the dataset.“The dataset store the image's geometry as well as phrases from the URL, the image's URL and alt text, the anchor text, and words that appear near the anchor text. In dataset, class labels (Binary Classification): "ad" and "nonad". The exciting part about this data is that as part of some pre-processing operation, someone could want to filter the WebPages for unnecessary ads (useful for later website categorization, for example.). Data mining is used to detect whether an image is an advertisement ("ad") ("nonad"). Because I had not provided any exact training/test split, I opted to adopt an acceptable way of measuring performance. After completing feature reduction to reduce the number of features, I used several data mining methods for prediction.

**Keywords–** Different Data Mining Algorithms, Chi-Square Method, Evaluating measures.