



CLASSIFICATION, SUMMARIZATION AND EXTRACTION OF DATA THROUGH ARTIFICIAL INTELLIGENCE

Sanjeet Banerji

EVP & Head, Artificial Intelligence & Cognitive Sciences

Document classification is one of the most important and fundamental feature of synopsis.

It uses a number of techniques like naive bayes distribution, random forest and in many situations we use an ensemble of multiple techniques to arrive at the right classification.

Classification refers to identifying the content of document and assigning it to a level where it belongs. One of the examples where classification can be very effectively used is claims classifications.

You get a number of claims for different kinds of diseases and the use of classification in this could help identify the nature of the claim and therefore put it in the right views for the right kind of processing for those claims.

Another very important feature of synopsis is summarization of documents.

Summarization is shortening

and extracting key information out of a long data or a long document.

Synopsis uses a number of techniques to summarize data it has. It provides a way to summarize data in form of percentage of typical documents it will create different kinds of summary of different forms and size. Summarization of medical records helps in identifying what kind of disease was reported where was it reported from and the frequency and the pattern in disease in claims.

Other features of synopsis include are things like topic models where it extracts key features or key topics that the document or a particular paragraph is talking about.

For example if you receive a lot of emails from customer care and you would like to find out what are these emails mostly about then topic features or topic models come to your help and they extract out the vast amount of such emails what are the key topics

that your customers are talking about which gives you an insight into how your business is going what your customers are talking about and what did they expect from you.

Synopsis also includes a number of other features like entity extractions which essentially means extraction of names, organizations, locations from a document. Location gives you geospatial information about the event or any incident or any claim that might be happening to in that location.

It helps you analyse the geospatial spread of your claims or any incidents that you might want to be interested in finding. Synopsis can ingest data from large number of sources that includes other databases which includes social media it includes URLs, shared photos and so on so forth. It creates its own repositories and you can always search the repository using semantic search or keyword search press.