A REVIEW OF CURRENT LEGAL AI DOCUMENT SEARCH AND ASSEMBLY SYSTEMS

July 9, 2016

Resilience Partners Ltd

Mo Afshar

Contents

1 INTRODUCTION

According to Blair Janis a faculty member at the Brigham Young University Law School teaching legal technology, the technology around the legal world advances at an exponential rate, the technology within the legal world, especially as it relates to lawyering (i.e., providing legal services as opposed to running a law business), is much slower. The main reason for this is that lawyers are generally risk averse. The reason behind this is lawyers need to be able to provide and obtain to a level of certainty that the service they provide will be successful.

As technology advances the law industry is required to have a high level of analysis to review and research within the said field so that they are able to implement it in the least risk oriented way. This paper will be a simple review of the current Artificial Intelligence legal systems, how and where they are implemented and how it can help the law scene to bring advantages to both the lawyers and the clients.

2   Artificial Intelligence in Law

Artificial Intelligence (AI) is usually defined as the science of making computers do things that require intelligence when done by humans or the power of a machine to copy intelligent human behaviour. The science revolves around learning, reasoning, problem-solving, perception and language-understanding.

2.1   AI in writing

AI based systems are already making an advance in industries such as journalism. An AI system called the Wordsmith, spots patterns and trends in raw data and then is able to translate those patterns found into natural language. The same or similar system can be used to produce legal documents, carrying out the tasks by possibly junior associates.

Electronic discovery (e-discovery) is an area where there are numerous examples of artificial intelligence used to reproduce the work of lawyers. It is essentially "to provide intelligent algorithms to find information based on concepts and key words agreed upon by the parties to the litigation". There is higher level of AI in e-discovery called the predictive coding, where a machine learns based on watching human behaviors and then by using the machine learning algorithms it will then apply what it learns. This sophisticated AI is then able to provide strategic guidance, this could be done by going through records of past cases and then a system can find a solution to would lead to a settlement. This has been tested by the London firm Hodge Jones and Allen to assess the viability of its personal injury caseload.

According to the firm "To develop the model, the law firm provided Andrew Chesher, professor of economics and economic measurement at University College London, with data about the outcomes of 600 cases concluded over 12 months. He used a combination of statistical techniques to examine the factors contributing to which cases were won or lost, the damages that were received by claimants in successful cases and the costs received by the firm." The factors examined were topics such as the nature and cause of the injury, the level of solicitor handling the case and the time between injury and instruction.

The results of the analysis were then used to produce further models to predict whether cases would be won or lost. The firm states that the model are being turned into Excel-based programmes which will aid their assessment team to help make initial assessment of the likelihood of a positive outcome for each case. However the firm stresses that it is not replacing

human skills, their senior partner says: "These models will not replace their experience and judgment, but will provide an additional aid to them in a world where it is no longer good enough to take a case on with a 50 per cent chance of success and where fees are restricted to a few hundred pounds."

2.2 SmartApps (Expert Systems)

Another form of AI for the legal industry is call the SmartApps, by having smart algorithms and more powerful computer processing, SmartApps address the legal challenges to provide precise and immediate answers to very specific questions such as if a payment can be made under the Foreign Corrupt Practices Act? Or what is my level of risk in making a transactions?

SmartApps are highly useful in regulated markets. They provide immediate answers to complex regulatory questions and complex international rules and restrictions, where a failure within these rules can cause very negative consequences and also take a lot of time to resource. Therefore, SmartApps free up the lawyers time to deal with more complex bespoke legal tasks. According to www.legalweek.com an example of this SmartApp is "Foley and Lardners FCPA Appv. This mobile app allows their clients sales teams to self-serve payment questions anywhere, anytime round the globe. It has provided Foley access to new clients deeper into the Fortune 500, extending their client base beyond what they were handling on a one-to-one human answer basis".

Another example by the USAs largest HR law firm is the self-service SmarAPp platform for HR questions. According to the firm "n this platform there are a series of Smart Apps including an Independent Contractor App that assesses the users employment status in 51 jurisdictions by assessing some 1,400 cases and 80 weighting factors."

Furthermore, ROSS is an artificially intelligent attorney which helps power through legal research. ROSS is built upon Watson, IBM's cognitive computers. Watson is able to mine facts and conclusions from over a billion of text documents a second. The link provided introduces Watson's technology and how it can be applied to different fields of work:
`https://www.youtube.com/watch?v=Y_cqBP08yuA`

3   CONCEPT CLUSTERING

Conceptual clustering is a machine learning paradigm for unsupervised classification developed mainly during the 1980s. Conceptual clustering is closely related to data clustering however conceptual clustering is also inherent to the description language which is available to the learner. It essentially aims to efficiently cluster and explain the data provided.

Cluster analysis divides data into groups that are meaningful or useful. Depending the goal the cluster will capture the data differently, if meaningful is the goal the clusters should capture the natural structure. Clustering for understanding plan an important role in how people analyse and describe the world. In the context of understanding data clusters are potential classes of techniques for automatically finding classes, classes being meaningful group of objects that share common characteristics. For example in information retrieval, the world wide web consists of billion of web pages and the results of a query to a search engine can result in thousands and millions of pages. Clustering can e used to group these search results into smaller and more meaningful clusters when a query is searched. Similar idea can be used in documents in legal systems where a subset of meaningful documents are only gathered from a cluster of thousands.

So in principle cluster analysis groups data objects based only on information found in the data that describes the objects and their relationship. The goal is that the objects within a group be similar to one another and different from the objects in other groups.

3.1   Different Types of clusterings

- A partition clustering is simple a division of the et of data objects into non-overlapping clusters such that each data object is exactly one subset

- A hierarchical clustering is when we permit cluster to have sub clusters, which is a set of nested cluster that organized as a tree.

- An overlapping or none-exclusive clustering is used to reflect the fact an object can belong to more than one group at the same time.

- A fuzzy clustering is when every object belongs to every cluster with a membership that either it certainly belongs or it certainly doesn't belong.

- A complete clustering assigns every object to a cluster, whereas a partial clustering does not.

3.2   Different types of clusters

- A well-separated cluster is a set of objects in which each object is closer to every other object in the cluster than to any object not in the cluster.

- A prototype based is a set of objects in which each object is closer to the prototype that defines the cluster than to the prototype of any other cluster.

- A density based cluster is a dense region of objects that is surrounded by a region of low density.

3.3   Popular techniques in cluster analysis

- K-Means Clustering is a method of vector quantization, k-means clustering aims to partition n observations into k clusters which each observation belongs to the cluster with the nearest mean.

- DBSCAN is a density based clustering algorithm that produces a partition clustering, in which the number of clusters is automatically determined by the algorithm.

- Agglomerative clustering is a "bottom up" approach: each observation starts in its own cluster, and pairs of cluster are merged as one move up the hierarchy.

- Divisive clustering is a "top down" approach: all observations start in one cluster, and splits are performed recursively as one moves down the hierarchy.

4   DOCUMENT ASSEMBLY SYSTEM

Document Assembly system is a software program integrated with one or more template documents and a text editor. To create a document from a template, the user answers a list of questions, based upon responses the program inserts the relevant information into the document, selects clauses from various alternatives and generates a first draft for a review. This process has been used for over 20 years where originally it was aimed to complete and automate repetitive processes that did not require much legal expertise.

The primary benefit of a sophisticated document assembly system is the speed with which high quality drafts of complex documents can be created. This brings significant advantage for clients such as banks, since they deal on a frequent occasions and speed is vital to their process getting their deals or not and additionally law firms would be able to deal with more complex matter more frequently. The reduction in the effort and costs associated with document drafting can drastically increase profits.

Furthermore, in Richard usskind's book 'The End of Lawyers' looks at the use of document automation software that enables clients to generate employment contracts. Additionally, document modeling looks at the inherent structure in documents. Document modeling is able to look at the structures and patterns of the written work and breaks it down into different options or branches. It then labels the branches and the results, without this it is hard to get the full value from a document automation.

REFERENCES

Blair Janis *How Technology is changing practice law* at:<`http://www.americanbar.org/publications/gp_solo/2014/may_june/how_technology_changing_practice_law.html`> [Accessed July 2016]

Michael Odell 2014. *Humans need not apply.* [online] Available at: <`http://raconteur.net/business/humans-needs-not-apply`> [Accessed July 2016]

Michael Cross 2015. *time for technology to take over.* [online] Available at: <`http://raconteur.net/business/time-for-technology-to-take-over`> [Accessed July 2016]

Sarah Whitten 2016. *the future of the legal system* [online] Available at: <`http://www.cnbc.com/2016/05/12/the-future-of-the-legal-system-artificial-intelligence.html`> [Accessed July 2016]

Greg Wildisen. *Is artificial intelligence the key to unlocking innovation in your law firm?* [online] Available at: <`http://www.legalweek.com/sites/legalweek/2015/11/12/is-artificial-intelligence-the-key-to-unlocking-innovation-in-your-law-firm/?slreturn=20160608133942`> [Accessed July 2016]

Wiki *Conceptual Clustering* [online] Available at: <`https://en.wikipedia.org/wiki/Conceptual_clustering`> [Accessed July 2016]

PDF Cluster Analysis, basic concepts and algorithoms. [online] Available at: `https://www-users.cs.umn.edu/~kumar/dmbook/ch8.pdf`> [Accessed July 2016]

Hierarchical clustering. [online] Available at: `https://en.wikipedia.org/wiki/Hierarchical_clustering`> [Accessed July 2016]

Document Automation. [online] Available at: `https://en.wikipedia.org/wiki/Document_` `automation>` [Accessed July 2016]