Using KOS for Artificial Intelligence Applications

Jay Ven Eman, PhD, CEO
Access Innovations, Inc.
j_ven_eman@accessinn.com

AIMS

There are many algorithms, which are classified as Artificial Intelligence (AI). Most AI systems use several. Some are based on absolutes and others on inference and statistics.

This paper reviews the current market conditions for AI, economic considerations, and technical barriers. Insuring replicable, reproducible accuracy in AI tagging and driving of business decisions is crucial. Our findings are that KOS driven AI is less expensive to implement and more accurate. (2) (3)

Machine learning and artificial intelligence approaches to automatic indexing and other aspects of content enrichment have tremendous potential, but there are significant barriers to successful implementations. The economics of these systems are not now generally affordable, which will indefinitely delay widespread adoption. (1)

The algorithms and theories used to support Artificial intelligence seem nearly endless. Some are absolute and replicable from one corpus to another providing a firm platform for reproducibility and commercial applications. For example, entity extraction and identifications (people, places, things), salience via weighting, syntactic analysis (parsing), semantic analysis, pragmatic analysis, grammar, lemmatization – stemming, morphological analysis and applications, lexical variations – synonyms, part-of-speech tagging, sentence boundary detection, punctuation marks as guides to conceptual sections, abbreviations, terminology extraction, term weighting, co-occurrence*, rules to increase accuracy, word parsing, phrase parsing, rule bases, concept extraction, spell checking etc.

Others are more statistical and supported by modeling and training. These may be subject to bias and require strict attention to the underlying training sets and models to insure replicable results and repeatable reproducible application. These applications include neural nets, Bayesian statistics, vector analysis, inference engines, co-occurrence*. They are usually more expensive to implement and maintain. * Co-occurrence can be more than one thing; counts of occurrences, parallel reference, in same article.

MAIN FINDINGS

There are a lot of risks in implementation of artificial intelligence including unethical use … insufficient learning … incomplete data… inaccurate data … unsecured data, regulatory noncompliance … unrepresentative data … biased models … discriminatory outcomes … model

instability … performance degradation … implementation errors … poor design, insufficient training … technology malfunction … performance issues … human machine interface failures.

Significant costs are involved in just the training and maintaining systems that chronically under perform and fail to scale. Cost and performance data will be characterized and presented. Machine learning and artificial intelligence projects are not for the faint of heart, nor for those with small budgets. Key cost elements are identified along with approaches to estimating costs based on actual and reported cases.

METHODS

This paper will cover these areas and then outline a parallel use case outlining the costs of implementation for a statistical system versus a rule-based system.

REFERENCES

(1) https://www.topbots.com/essential-landscape-overview-enterprise-artificial-intelligence/

(2) Here and now: Bringing AI into the enterprise *Kristian Hammond (Northwestern Computer Science)* 9:00am–12:30pm Tuesday, June 27, 2017 https://conferences.oreilly.com/artificial-intelligence/ai-ny-2017/public/schedule/detail/59188

(3) https://www.kungfu.ai/the-essential-landscape-of-enterprise-ai-companies-topbots/