An AI based Text Analytics Suite
Businesses face challenges with deciphering text data in various complex formats thereby limiting the ability to derive insights. tex.ai, a SaaS based solution from Indium will help solve text data problems, by providing faster and scalable solutions to see impressive results such as saving time by summarizing the document, lowering costs by removing manual extraction, reducing search complexity by classifying documents, improving customer experience and many more.

A team of in-house NLP, Ai / DL experts have been working with several clients and have working prototypes for a variety of text analytics services.

**Offerings**

*From unstructured text to insights in a jiffy!*

**Classification**
- Identify the category for the text descriptions
- [customer type as irate, satisfied, indifferent]

**Summarization**
- Extract key phrases, topics from reviews and documents; find the topic clusters at individual and aggregate levels
- [topics -> words, price 23% (very costly, low priced), quality 54%(durable in tough, lasts for long...), delivery 45% (was quick, within 2 days...)]

**Extraction**
- Extract structured, unstructured tables and table look-alikes from documents of different formats
- [image of tables in pdf to neatly formatted tables in csv]
Value Proposition

- Automate text extraction from PDFs, images and websites to structure the unstructured data.
- Get a bite-sized summary of huge TI; DR documents such as articles, reviews, tweets, comments, legislation and more.
- Assign a category to each document (books, articles, journals, reviews, tweets, comments, legislation).

Technologies

Text Extraction
- Libraries Used: Tabula, Camelot, Tensorflow, Keras, OpenCV, Pytesseract, BeautifulSoup, Scrapy, Selenium

Text Summarization
- Libraries Used: TNLTK, re, Keras, scikit-learn, Pycrsuite, ELMo, NetworkX, spaCy, Tensorflow
- Algorithms: Dependency Parsing, POS based Grammar Chunking, TF-IDF, Non-Negative Matrix Factorization, k-Means, tSNE, Network Analysis, Conditional Random Field, LSTM

Text Classification
- Libraries Used: Genism, SpaCy, scikit-learn, Tensorflow, ELMo
- Algorithms: TF-IDF, Word2Vec, Naive Bayes, SVM, Neural Networks

Domains

- Financial Research Firms
- Rating Firms
- Content Aggregators
- E-commerce Aggregators
- BFSI
- Legal Firms