

# Text Mining for a fastest growing Indian Digital Business - Case Study



## Client Overview

- The client is a leader in the \$200 B local services market in India and for NRI's in USA, Canada, UK and UAE
- The client's platform helps users to minimize the time in finding the right service provider, to reduce the cost of the service and to minimize the hassle of dealing with service providers
- The client's platform connects 25 million users with 5 million businesses each month across different local service categories each month



## Business Requirements

- The client runs a lead generation service in the local services space. When a user wants to request a service, they are required to fill up a Lead capture form, which requested details relating to the service
- In addition to basic fields, there is also a provision for the user to enter free text pertaining to the service required
- The client wanted a solution to glean useful insights from the text and reduce operational overhead



## Key Challenges

- The client received 1 million rows of free text per month related to the service requests in their online platform
- The challenge was to identify the groups of most commonly repeated words with higher accuracy



## Our Approach

- We analyzed the raw lead capture form free-text data and chose TF-IDF & Support Vector Machines as the appropriate text mining algorithms
- The data required structuring into a format ingestible by text mining algorithms. This required significant data wrangling, which was carried out using R & Python



## Our Solutions

- We applied text mining techniques & algorithms such as TF-IDF & Support Vector Machines, and were able to isolate the most important words being mentioned repeatedly
- This led to the product team incorporating frequently requested service options and including them in the lead capture form



## Results

- This continuous analysis of the text corpus and gleaning of insights led to an improvement in overall conversion rates yielded from the Lead capture form
- The solution increased the overall operational efficiency; as efforts required by the team which manually analyzed the free-text content was greatly reduced



INDIA  
Chennai  
+91 44 6606 9100  
Bengaluru  
+91 80 4645 7777  
Mumbai  
+91 022 6215 4028

USA  
Cupertino | Princeton  
Toll-free: 1 888 207 5969  
SINGAPORE  
+65 9630 7959

UK  
London  
+44 773 653 9098

General Inquiries: [info@indiumsoftware.com](mailto:info@indiumsoftware.com) | Sales Inquiries: [sales@indiumsoftware.com](mailto:sales@indiumsoftware.com)

[www.indiumsoftware.com](http://www.indiumsoftware.com) | ©Indium Software