Digital Media Analytics: Customer Segmentation & Churn Analytics

How can we transform your business through digital analytics

Customer Segmentation & Churn analytics

Setting the Context: Trends in Media & Entertainment World

Table: India Growth Story of Television vis-à-vis Digital Media

	CY2016	CY2017	CY2018E	CY2020E	CAGR 2016 - 20
Television	594	660	734	862	9.8%
Digital Media	92	119	151	224	24.9%

Data Source: https://www.ey.com/in/en/industries/media---entertainment/ey-re-imagining-indias-me-sector

- Indians' appetite for binge-watching online videos is set to make the country among the top ten OTT (over-the-top) video markets in the world in four years*
- Recent report** on M& E industry reveals that there are around 1-1.5 million "digital only consumers" in India today, who do not normally use traditional media
- This customer base (is expected) to grow to around 4 million by 2020 and generate significant digital subscription revenues for the media and entertainment (M&E) sector
- Tactical digital consumers — who consume both paid television (cable and DTH) and have at least one OTT (over the top) subscription (like Netflix, Hotstar or Amazon Prime Video), on the back of digital and micro-payment systems being rolled out in the country, will reach as high as 20 million households by 2020, from 6 million now
- By 2020, the segment of consumers that consume traditional media (either pay or free) and free OTT content is expected to cross 500 million. Overall, the M&E industry is poised to touch $31 billion by 2020

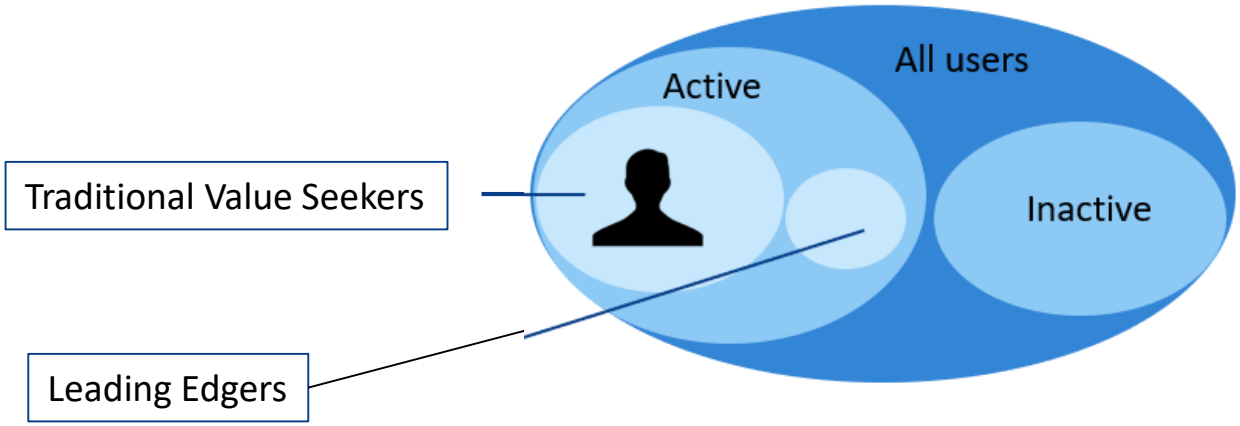
*Source: PwC India

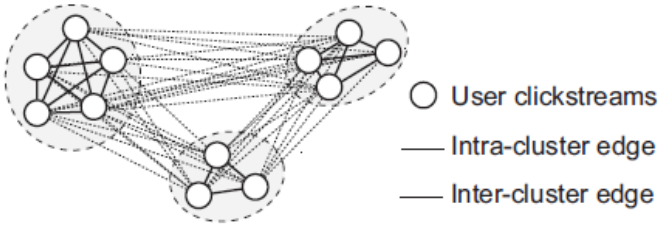
**Source: 'Re-imagining India's M&E sector' published by EY & FICCI in March 2018

Solution # 1: Unsupervised Clickstream Clustering
Segmentation & Developing Buyer Persona

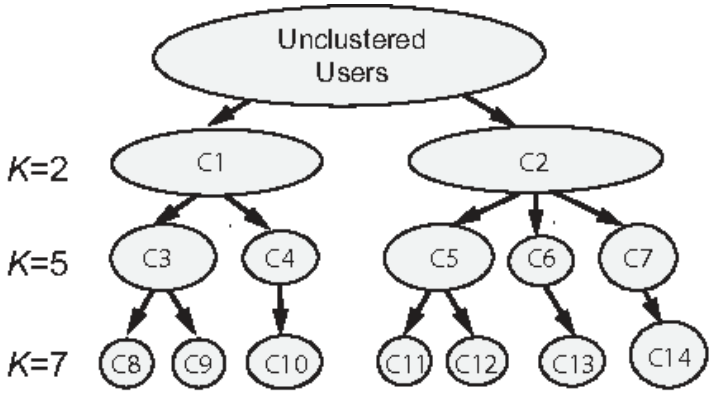
Key Intuitions

- ✓ Users naturally form clusters
- ✓ More fine-grained user clusters are hidden within big clusters





Clickstream Similarity Graph



Hierarchical Structure of Behavioral Clusters

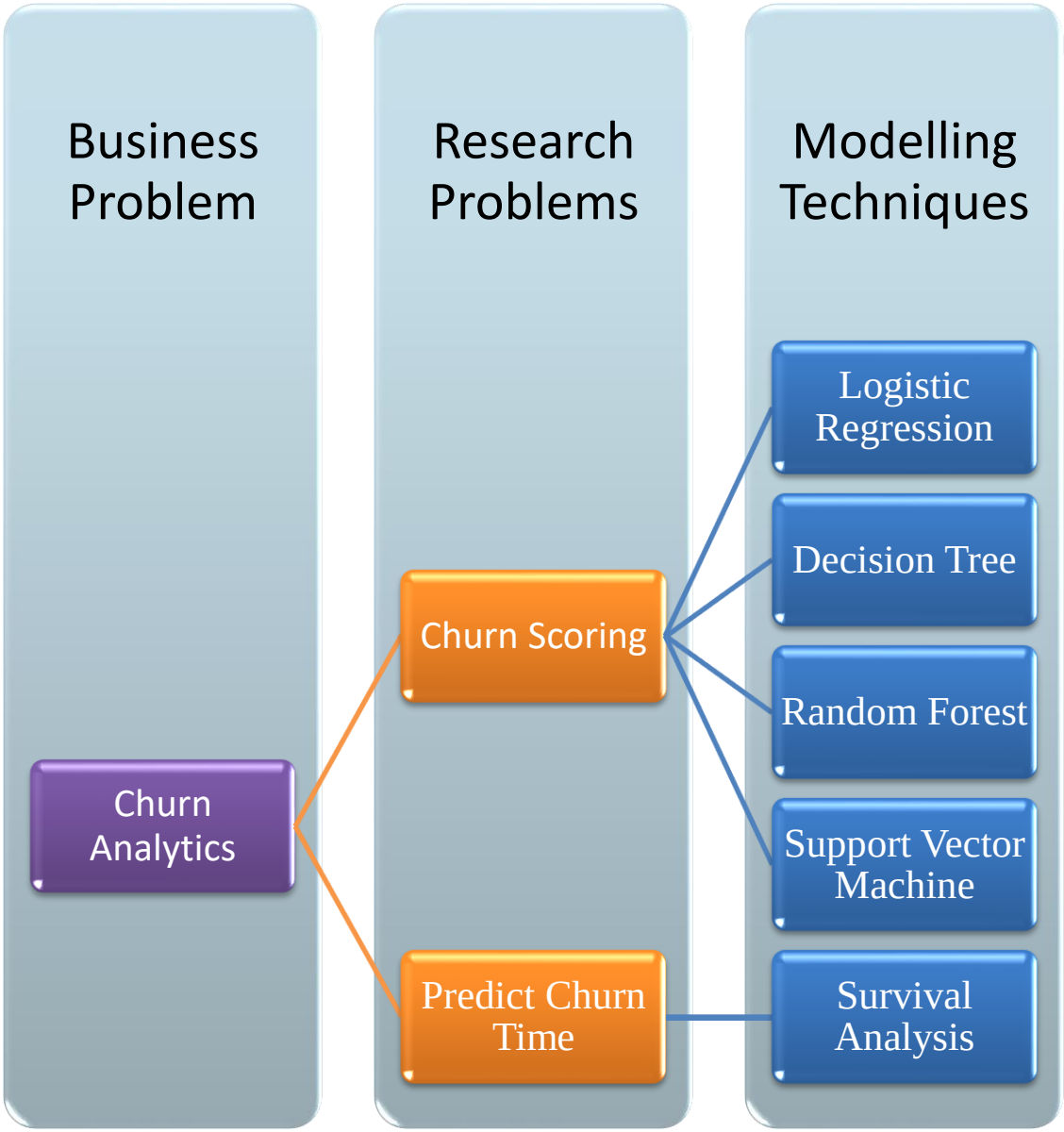
The Outcome

- ➢ Identify natural clusters of user behavior based on clickstreams
- ➢ Extract semantic meanings for captured behaviors
- ➢ Scalable for large online services

Solution # 2: Churn Analytics

Churn Analytics Objectives	Which customers are going to churn?
	When the customer is going to churn?
	Why are the customers leaving? What factors are critical for a customer to churn?
	How does a churner profile look like? What makes our churner different?

Proposed Solution Architecture

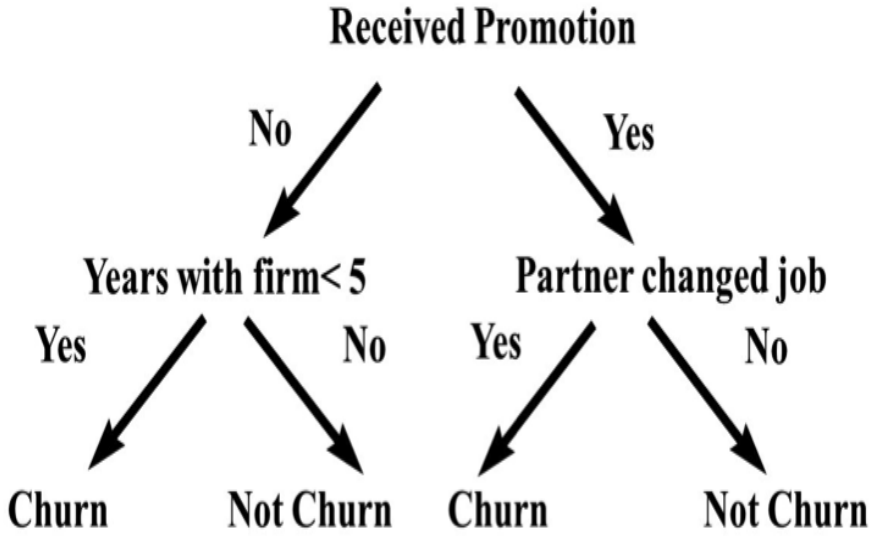


Note: One or more than one algorithms would be used depending on the nature of the problem/ data

Churn Scoring Algorithms

1. Logistic Regression: Logistic regression is used to describe data and to explain the relationship between one dependent binary variable (active/ inactive customers) and one or more nominal, ordinal, interval or ratio-level independent variables.

2. Decision Tree: A decision tree is a decision support tool that uses a tree-like graph or model of decisions (churn) and their possible reasons.



3. Random Forest: Random forests or random decision forests are an ensemble learning method for classification, regression and other tasks, that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees.

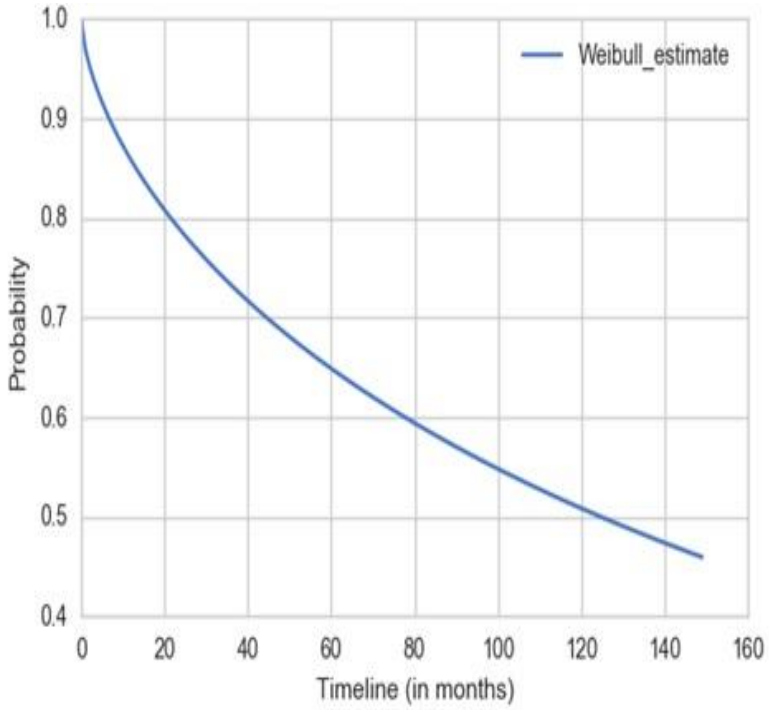
4. Support Vector Machine: each data item is plotted as a point in n-dimensional space (where n is number of features) with the value of each feature being the value of a coordinate. Then, the classification (churn/ not churn) is performed by finding the hyper-plane that differentiate the two classes very well.

Churn Time Prediction Algorithm: Survival Analysis

Survival Analysis is a technique where the outcome variable is the time until the occurrence of an event of interest (churn). Hazard Function or hazard rate h(t) is the conditional probability that the event (churn) will occur within a small-time interval after t, given that it has not occurred until t.

Survival Analysis functions:

1. Exponential: Parametric estimate where hazard is constant over time
2. Weibull: Parametric estimate where hazard function is increasing or decreasing over time
3. Cox Regression: Semi-parametric estimate for proportional hazards regression
4. Kaplan-Meier: Non-parametric estimate commonly used to compare two populations



Key Benefits of Churn Analytics

1. Prevention is better than cure: Customer churn reduces profitability through revenue loss

2. Churn results in greater marketing and re-acquisition costs: The customer acquisition cost is higher than the cost of retaining an existing customer… sometimes by as much as 15 times more expensive

3. The probability of selling to an existing customer is a lot higher than to a new prospect

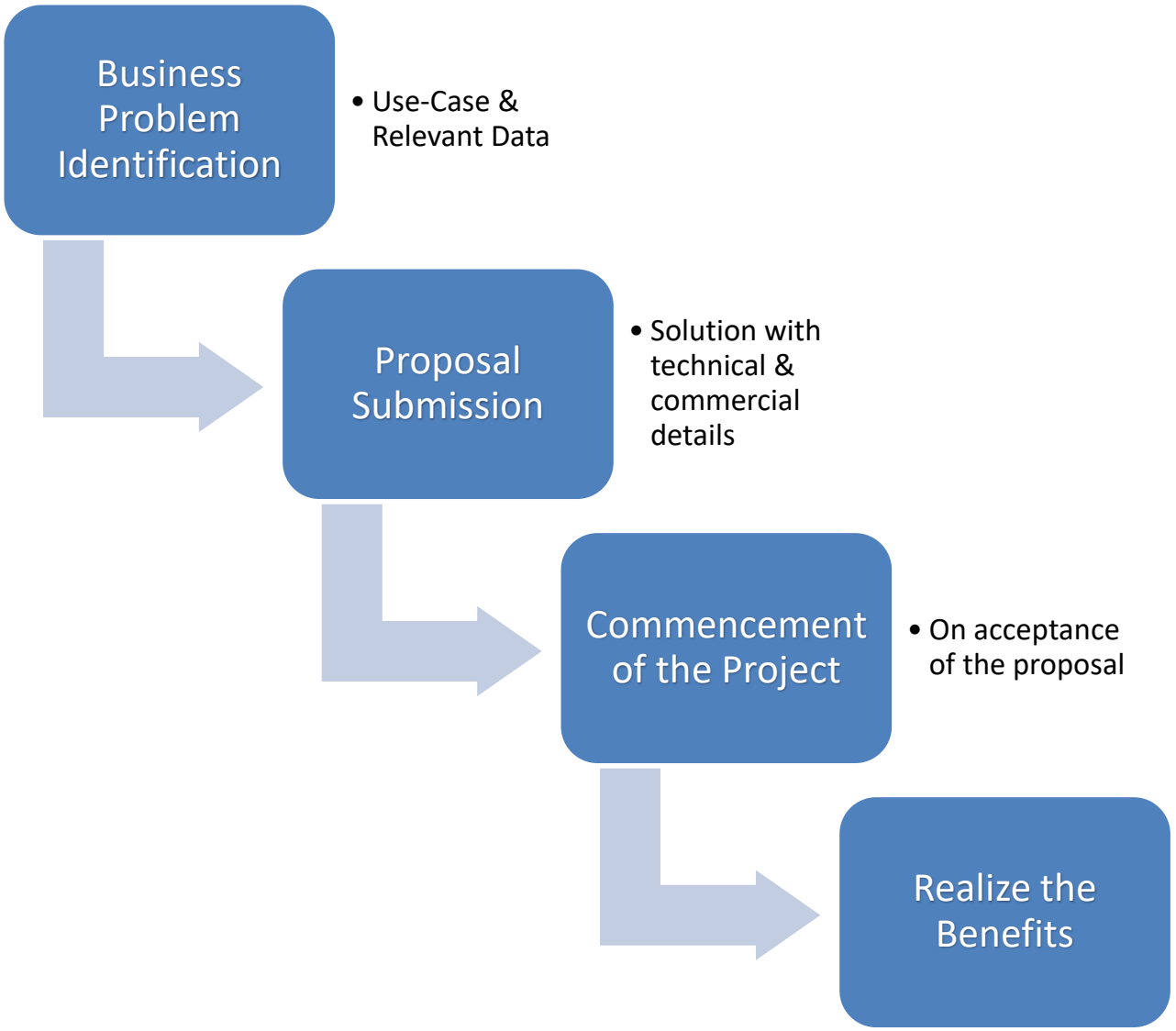
4. Decrease the likelihood that competitors will lure existing customers

5. Proactively detect customer value loss and take measures

Agenda # 3

How to Get Started

The Road Ahead



Disrupt or Get Disrupted! Act Now

Digital Media Analytics is fundamentally Disrupting Traditional Business Models

Digital Media Analytics offers NEW Challenges & Opportunities

Time to act is NOW due to imminent huge growth wave

QuantFarm has that special X-Factor with the right expertise and experience to help you embark on Supply Chain Analytics

Thank You