Integrated Media Planning: Business Analytics and Model building

SANGEETA SHAH BHARADWAJ

SSBHARADWAJ@MDI.AC.IN

Case Study: Toyota Camry

Today's consumers are exposed to an expanding, fragmented array of marketing touch points across media and sales channels.

Imagine that while viewing a TV spot for a Toyota Camry, a consumer say Rita uses her mobile device to Google "sedans. " Up pops a paid search link for Camry, as well as car reviews

She clicks through to Car and Driver's website to read some reviews, and while perusing, she notices a display ad from a local dealership but doesn't click on it

One review contains a link to YouTube videos people have made about their Camrys. On YouTube she also watches Toyota's clever "Camry Reinvented" Super Bowl ad from eight months earlier

During her commute to work that week she sees a Toyota billboard she hadn't noticed before and then RELATED VIDEO receives a direct-mail piece from the company offering a time limited deal

She visits local dealerships' websites, including those promoted on Car and Driver and in the direct-mail piece, and at last heads to a dealer, where she test-drives the car and buys it.

How many channels has she touched?

- TV
- Youtube, website, google ads
- Billboards
- Radio
- Direct mails
- Print media ad

It becomes impertinent to talk about Various Media Platforms

- TV
- Radio
- Mobile Apps or Alerts
- Paid Search
- Organic Social Media Content
- Website Features
- Influencer Campaigns
- External Press Coverage
- Paid Social Media
- Media: YouTube videos, podcasts
- etc
- etc

What questions should Toyota's chief marketing officer should ask ?

- ??
- Customer persona
- Areas visiting
- Design Campaigns and place advertisement
- What is my target segment?
- Target location?
- Which advertising media? Garners maximum attention
- Preferred platforms for information consumptions?
- How do you arrive at right media mix plan? **Cost and benefit??**

What questions should Toyota's chief marketing officer (digital Marketing) should ask ?

- ??

- How did this combination of ad exposures interact to influence this consumer?

- Is Toyota investing the right amounts at the right points in the customer-decision journey to spark her to action?

Another case

How do ads work across media and sales channels, following is an example

- TV ate up 85% of the budget in one new-product campaign, whereas YouTube ads—a 6% slice of the budget
- However You Tube ads were nearly twice as effective at prompting online searches that led to purchases
- And search ads, at 4% of the company's total advertising budget, generated 25% of sales.
- Armed with those rich findings and the latest predictive analytics, the company reallocated its ad dollars, realizing a 9% lift in sales without spending a penny more on advertising

What is different here?

- ??

So what is different in this case?

??

- Capture the data
- Use predictive analytics
- In this case improved sales by just reallocation without spending extra money.

Hence it is important to study analytics arounds advertising spend , media mix and timings of ads, impact on sales
How Data Analytics is helping ?

Media Mix Analytics

- What are the different ways of setting advertising budget?
  - Media decisions (what mix?)
- What are the methods that involve measuring advertising effectiveness?
  - Advertising effectiveness (can normally be done by conducting experiments in controlled advertisement or otherwise)

Media Mix Analytics

- What are the different ways of setting advertising budget?
  - Media decisions (what mix?)
- ~~What are the methods that involve measuring advertising effectiveness?~~
  - ~~Advertising effectiveness (can normally be done by conducting experiments in controlled advertisement or otherwise)~~

Media Mix Budget Decisions

- Advertising expenditure is percentage of sales, how do you decide?

Media Mix Budget Decisions

- Advertising expenditure is percentage of sales, how do you decide?
  - Delphi techniques
  - Benchmarking: Advertising investment varies by industry(B2B: Product and Services, B2C: Product and Services)
  - Competitors analysis (what is competition doing, what is the spending etc)

What else is important?

- Sensitivity of sales or other variables to different input variables
- What does that mean?
- One example is…

How sensitive is demand to advertising?

Advertising elasticity of Demand= Proportionate change in demand for the product/ proportionate change in advertising expenditure

If Advertising elasticity of demand (AED) is more than 1 , the product is advertising elastic

If Advertising elasticity(AED) is less than 1, the product is advertising inelastic

Is this all you want to know?

- ??
- Sensitivity of corona cases on Pay per click?
- Sensitivity of GDP on Sales?
- .
- .
- .
- .

Decomposition One way of knowing the ad reach is

- This method allows the advertiser to obtain a measure of the reach of the advertisement.

Total Audience	400 million	
Total cost of Ad	2,000,000	
Profit from each unit sold	Rs 1,000	
	Percentage	**Percentage of audience**
Total market	40%	40%
Exposed to ad	60%	24%
Pay attention to ad	40%	9.6%
Understand Ad	70%	7.72%
Intend to purchase	10%	0.67%
Actually purchase	20%	0.13%
Customers : 53,760	Profit: 53,760,000	ROI: 25.88%

If you are doing digital ads on social media how do you think these figures changing?

So what next?

- If you know reach of media let us say impact of media in garnering interest/ sales, what next??

- You may want to collect this information for all media points
- You will have to analyze how to allocate the budget to each media channel

Allocation of budget: Optimization problem

Objective is to find the promotion mix that maximizes total number of impressions (from all channels together)

The channels being considered are Direct marketing, Pay-per click(PPC) advertising and social media advertising

For direct marketing the available budget if Rs500, and the impression achieved per advertisement is 300

For PPC the available budget per advertisement is Rs 600 and the impression achieved per advertisement is 300

For social media the budget per advertisement is Rs 400 and the impression achieved per advertisement is 400

Cont.

The constraints are Direct marketing can not exceed 2,400 ads, PPC advertising can not exceed 1,600 advertisement and social media can not exceed 800advertisement per month

The objective function is to maximize total advertising impressions for the month promotion mix across the three channels

Another example of Media selection Models

- Show excel sheet

Profitability Model for Pay Per click(PPC)

- Estimated cost per click: Rs 1
- Estimated clicks per day: 10
- Conversion rate: 5%
- Avg profit per sale Rs 10

Monthly clicks	300 (assume 30 days in a month)
Click cost	Rs 300
Conversion per month	15
Profit	Rs 150
Total monthly Profit	-Rs150

So how do we conclude with these examples?

??

Marketing Analytics in Media mix involves three broad activities

- ???

- attribution, the process of quantifying the contribution of each element of advertising;
- optimization, or “war gaming” by using predictive analytics tools to run scenarios for business planning; and
- allocation, the real-time redistribution of resources across marketing activities according to optimization scenarios.

Attribution

To determine how your advertising activities interact to drive purchases

start by gathering data, What type of data and Where do you find that data?

market conditions

marketing actions

competitive activities

Individual swimlane without attribution

ADVERTISING MEDIUM	ESTIMATED RESULTING REVENUE
DISPLAY ADS	$40 MILLION
PAID SEARCH	$50 MILLION
SEARCH ENGINE OPTIMIZATION	$40 MILLION
E-MAIL MARKETING	$30 MILLION
TOTAL	**$160 MILLION**

the marketing team presented to finance some campaign results that had been generated using traditional analytics methods

Things quickly became awkward when finance pointed out that the business unit had generated only $110 million in revenue, $50 million short of the reported total.

The discrepancy arose because, leaders in each swim lane claimed the same bucket of revenue

Attribution Models

Management Slant

- At each stage in a consumer's journey toward purchase, different online channels feature most prominently.
- Existing credit-assignment methods, such as the last click, suffer from the problem of attribution—they do not take into account the impact of all those advertising formats that were visited by a consumer contemplating a purchase.
- Four rule-based models can be used for measuring the performance of an advertising campaign—the last-click, time-decay, uniformly distributed, and position-based models.
- Multichannel attribution models have evolved to reflect the growing complexity of attributing credit with each new advertisement format.

Need to study the following

- at what stage in a consumer's journey different online channels feature most prominently for an online business;
- the financial importance of these channels under last-click models;
- the effects of moving to rule-based multichannel attribution models—time decay, uniformly distributed, and position based—and statistics-based multiattribute models.

The study investigated

- whether multichannel attribution models give different channel valuations than last-click models;
- whether these channel valuations vary significantly among the various multichannel models;
- whether statistical multiattribute models have predictive validity.

Following scenarios are discussed

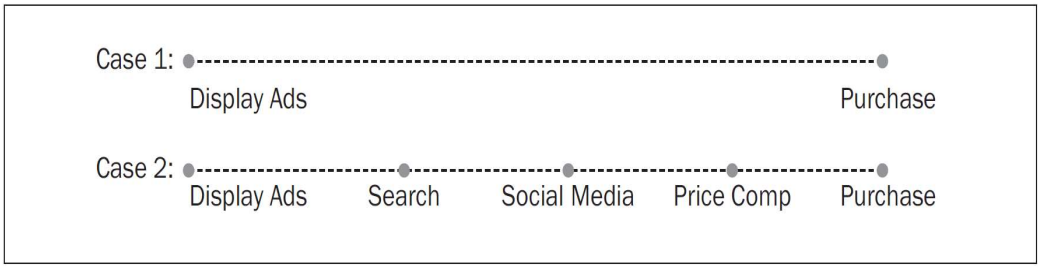
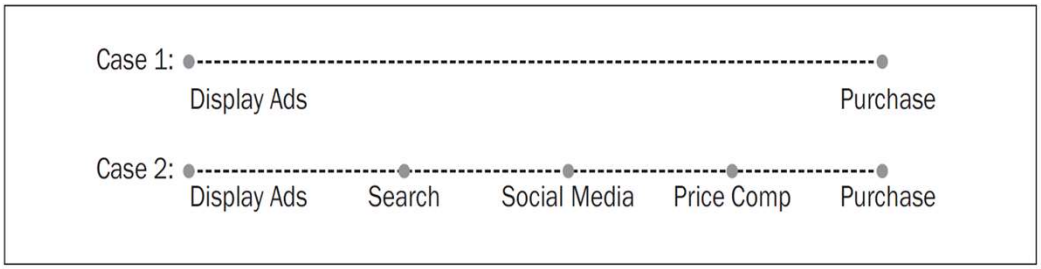


Figure 1 Conceptual Model of Multiplatform Advertising Strategy and Brand Perception

The Role of Online Sales Channels In the Customer Journey

A customer can take different choice actions at different stages of a purchase funnel. There can be three roles in a customer journey—introduction, assist, and conversion (Chandler-Pepelnjak, 2010;

The last click model



The last click model ascribes 100% credit to the last advertisement clicked o before a purchase conversion so in case 2 model would attribute the entire conversion to price comparison

Multichannel attribution Model

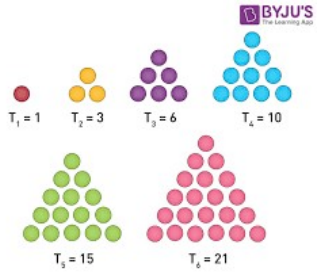
Model no.	Attribution model	Modeling approach
1	Time-decay model	Rule-based modeling assumptions
3	Uniformly distributed attribution	Rule-based modeling assumptions
2	Position-based model	Rule-based modeling assumptions
4	Statistics-based model	Cooperative game theory-based modeling assumptions

The time decay model

This rule-based attribution model follows the triangular numbers ratio of 1:3:6:10. In order to calculate the ratio breakdown for different-length journeys, therefore, the authors used the following formula:

$$\mathrm{Tn} = n\frac{(1+n)}{2} \quad (1)$$

In the above formula, the use of n represents what step number it is; "Tn" is the weighting given to it, and the 100 percent commission is divided up depending on the ratio. A three-step journey, for example, will be divided up under the ratio of 1:3:6, with 10 percent, 30 percent, and 60 percent attributed to the steps in chronological order. For each marketing tool, the model



Uniform Distribution Model

In a uniformly distributed multi-impression attribution model, the value of each conversion is distributed uniformly to all impressions. In Case 2, the uniformly distributed model would attribute 25 percent of the conversion to each of the four channels involved. Because the shares of each conversion are divided equally among all channels, the model does not consider where the touch points occur.

The Position Based Model

A popular position based model uses pareto distribution and hence places gigh values to first and last purchase (80% attribution and equally distribute rest 20% on all other channels

Results of descriptive statistics

TABLE 2 Different Online Marketing Tools and Revenue Generated under the Last-Click Method

Tool	Revenue (%)	Orders (%)	Average order value ($)
Organic search	63	67	106
Display	18	13	159
Paid search	11	10	116
Others	3	3	113
Price comparison	2	2	136
Retargeting	1	2	110
E-mail	1	1	112
Social media	1	2	48

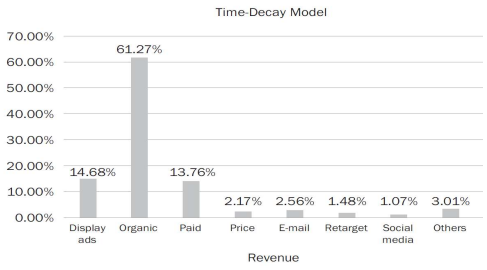
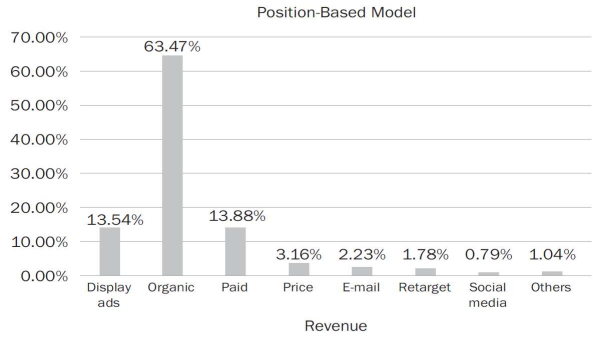
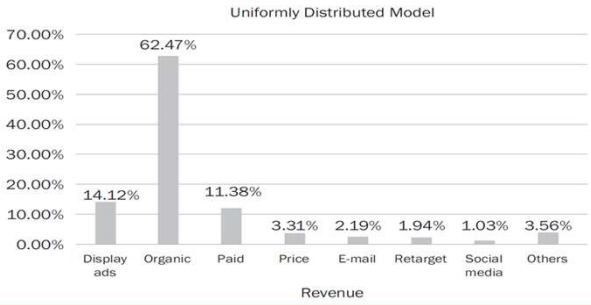
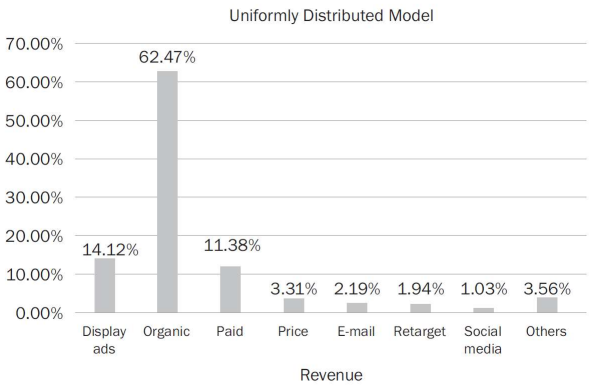


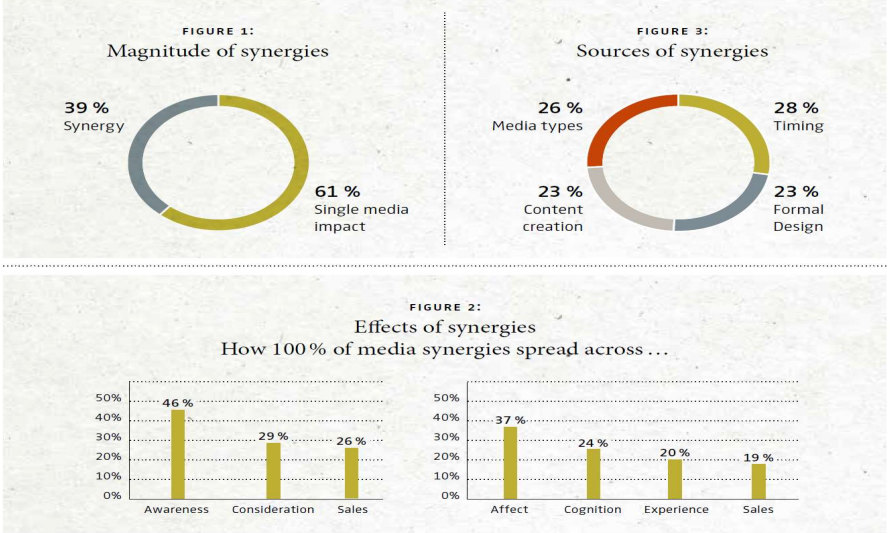
Figure 2 Online Revenue Generation According to Time-Decay Attribution Modeling







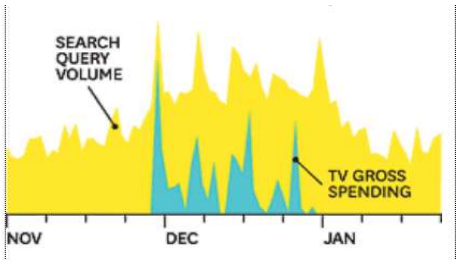
So what we are saying here is that there is synergy because of multimedia channels i.e. 1+1+1+1+1=5

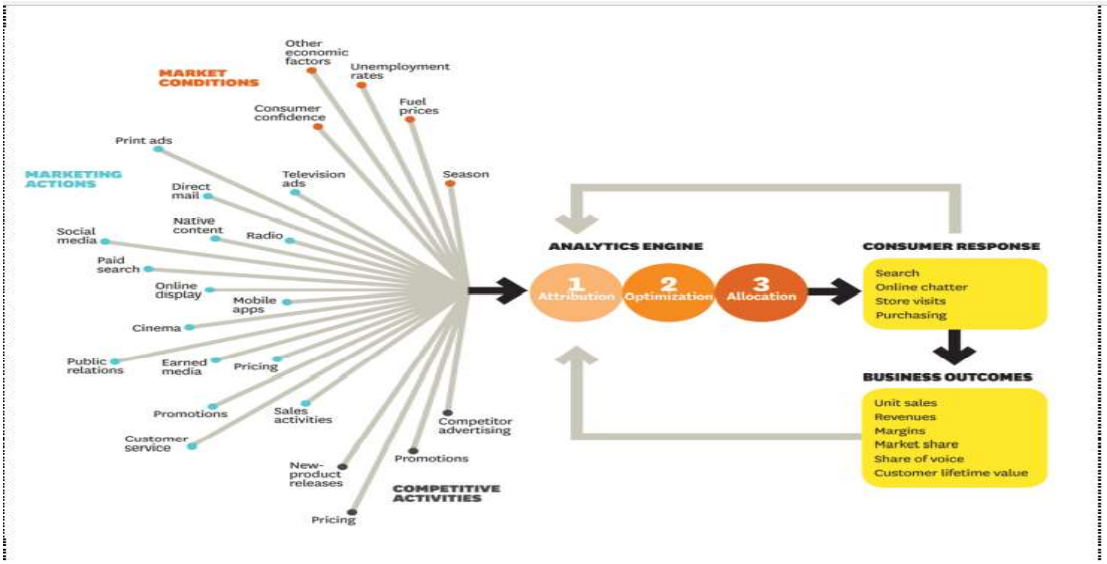


Advertising Analytics

- In actual analyses statistical models may account for hundreds or thousands of permutations of advertising and sales tactics, as well as exogeno variables such as geography, employment rates, pricing, season of the year, competitive offering and so on

- such analyses allows to instantly see how a ne TV ad affects consumers' online search patterns and then to change your keyword-search bidding strategy to buy up relevant words as the ad is running.



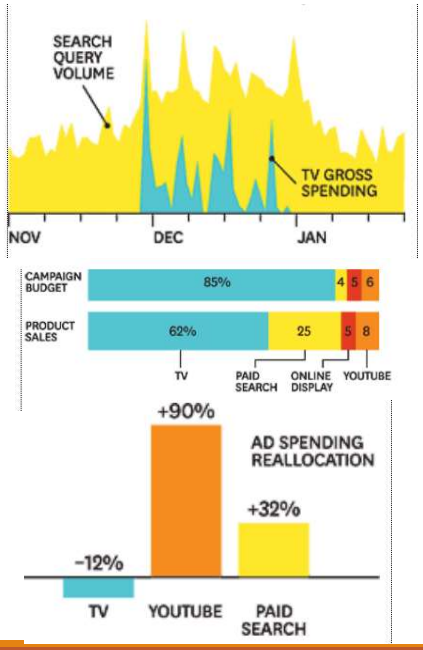


Optimization

- Once a marketer has quantified the relative contribution of each component of its marketing activities and the influence of important exogenous factors, war gaming is the next step
- It involves using predictive-analytics tools to run **scenarios for business planning**
- Maybe you want to know what will happen to your revenue if you cut outdoor display advertising for a certain product line by 10% in city A—or if you shift 15% of your product-related TV ad spending to online search and display

Remember Elasticity

- Working with data analyzed through the attribution process, you can assign an "elasticity" to every business driver, from TV advertising to search ads to fuel prices and local temperatures
- (Elasticity is the ratio of the percentage change in one variable to the percentage change in another.) Knowing the elasticities of your business drivers helps you predict how specific changes you make will influence particular outcomes
- If your TV ads' elasticity in relation to sales is .03, for example, doubling your TV ad budget will yield a 3% lift in sales, when all other variables remain constant
- In short, analytics modeling reveals how all driver elasticities interact to affect sales

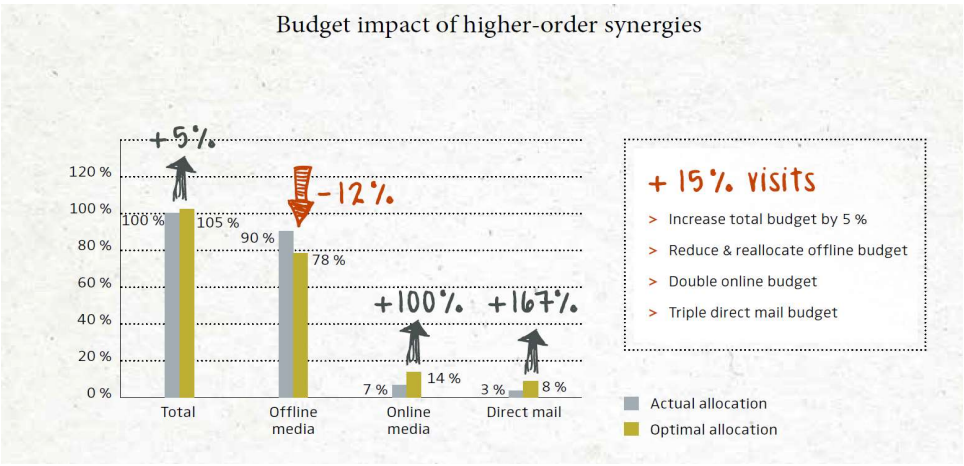


War Gaming

- War gaming uses the actual elasticities of your business drivers to run hundreds or thousands of scenarios within minutes
- In a typical war-gaming process, team members define marketing goals (such as a certain revenue target, share goal, or margin goal), often across multiple products and markets
- Crunching the vast database of driver elasticities, optimization software generates a set of most-likely scenarios along with marketing recommendations to achieve them
- The software also can test specific what-if scenarios: For instance, how will sales of our midsize pickup truck in Denver be affected if gas prices climb 5% and we launch a combined TV and online campaign promoting a $300 rebate?

Allocation

- Allocation involves putting the results of your attribution and war-gaming efforts into the market, measuring outcomes, validating models (that is, running in-market experiments to confirm the findings of an analysis), and making course correction
- Gone are the days of setting a marketing plan and letting it run its course—the so-called run-and-done approach.
- Advertising has become easier to transact, place, measure, and expand or kill.
- Marketers can now readily adjust or allocate advertising in different markets on a monthly, weekly, or daily basis—and, online, even from one fraction of a second to the next.



Can we summarize our learning here

- No one media is the best
- We need to swift through large set of data
- Analyze them in real time and take fast decisions

Integrated Media Planning

- Paramount to a successful **media** plan is the right use of online and offline data
  - to create consumer profiles
  - identify the best channels
  - Define time, space to drive exposure and action

Bindi Menon, interim chief marketing officer for Captain D's, LLC, said.

- "Keys to our success include reaching our consumers throughout their day
- delivering greater frequency to our most valuable consumers
- using data to drive **media** decisions
- The final key to success that we have found in an omni-channel environment is consistency. We strive to deliver relevant and fully **integrated** messaging to our consumers using complementary **media** channels for every promotion."

Define Audience?

- Define the Audience with Offline and Online Data and Insights
- Pairing what is known about consumers from offline data with online data provides a holistic view of a brand's and retailer's best consumer
- Offline data such as demographic, lifestyle and purchase habits can help identify high-value locations
- Online data captured from digital devices brings real-time consumer-centric insight that when paired with offline data, delivers a richer picture of the target audience and how, when and where to activate them.

Target Right Location

- Target the Right Location
- Today's consumer is on the go – not tethered to a single location – and receives promotions throughout the day from print sources, in-store, email, the internet and mobile
- This has led to an evolution in targeting to redefine geography around a true day-in-the-life of the consumer
- The more traditional retail trade area was based on the proximity of a retail store to a consumer's home
- Now the focus is on where a shopper "lives" throughout the day, leveraging additional data sets.

Media as Complementary, not Competitive

- multiple **media** touchpoints to inform and accelerate customers path to purchase
- uniformity in messaging and cross-channel optimization-- produces **media** efficiencies,
- Print and digital **media** do not compete for a consumer's attention, but rather complement each other as touchpoints throughout the course of the day.

Omnichannel Marketing

- omnichannel marketing offers the most benefits
- Omnichannel marketing uses different channels to seamlessly funnel customers all toward one single goal (say, to buy a book or set a consultation appointment)
- Marketers across industries agree that omnichannel is indeed the future of marketing

Questions?

Role of AL/ML/DL

AI: Definition

AI is defined as the assemblage of technological components that collect, process and act on data in ways that simulate human intelligence.

Like humans AI solutions can apply rules, learn overtime through execution of data and information (i.e., via ML)and adapt to changes in their environment

Key Components of AI Solution

Input Data

- **Historical**-e.g., Customer's past transaction, external credit rating information
- **Real time**-e.g., Beacons in stores, or tracking of online activities
- **Knowledge**-e.g., Past products recommendations which are accepted or rejected

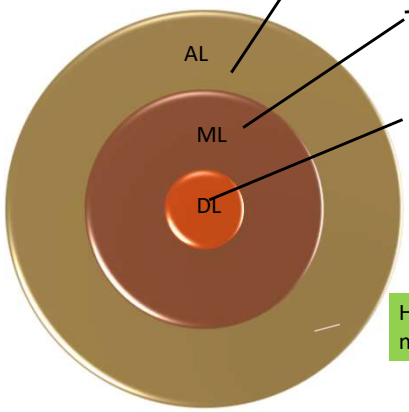
Processing Algorithm

- **Supervised learning** -e.g., Cancer detection
- **Unsupervised learning-** e.g., Identify items that are bought together, or to find related images in photo database
- **Reinforced learning**-e.g., Board game

Output Decision

- **Results**- e.g., Credit rating score
- **Selection of results-** e.g., Selection video that may violate YouTube's terms of service, and which need further analysis by the team of content curators
- **Action**- e.g., Self driving car has full autonomy to drive , steer or break

Categories within AI



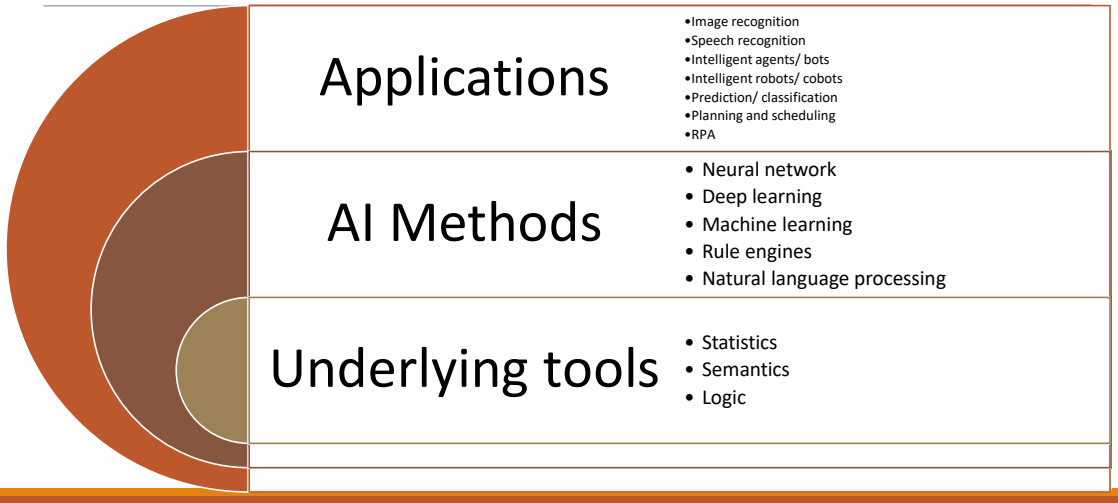
AL: algorithms that simulate human intelligence

ML: algorithms that can learn from data without relying on rule-based programming. Hence learn and improve(when more data is fed) without the intervention of human beings as how to do it

DL: Application of ML to complex problems, learn from mistakes and assess its own probability of reaching a correct result. With NN algorithms of DL additional data implies better predicting power

How can we learn from Data? And data can be numbers, text video etc etc

AI Tools, Methods, Applications



DL/ML/AI

In its most basic form, machine learning is synonymous with predictive analytics. Computers use models from data for which the results are known to predict the results from new data

Over the past several years, reinforcement learning, in which machine learning models are designed to maximize performance in support of a specified goal, has also become popular

Deep learning is used in image recognition and speech processing.

Unlike earlier forms of statistical analysis, the individual features of a deep learning model have little meaning to a human observer, sometimes to the point of being impossible to interpret

As a result, highly regulated industries like banking, insurance, and healthcare find deep learning models problematic, though researchers are attempting to make them more transparent

Key Enterprise AI trends

embedding AI into transactional systems

- For example, if a company wants to qualify and rank its sales leads by predicted likelihood of purchase (typically through some combination of machine learning and natural language processing), it has two choices; it can develop its own AI application to predict and score leads, and then try to integrate it with its customer relationship management (CRM) system, or it can buy the same capabilities from an established CRM vendor

democratization through automation

- The scarcity of trained personnel, including data scientists and AI engineers, has always been a substantial constraint on the use of AI.
- AI vendors offer automated machine learning, which makes it possible for less skilled analysts—sometimes called citizen data scientists—to do more sophisticated work
- Examples are data preparation, feature engineering or variable transformation, exploring different algorithms, selecting the best model, writing program code or APIs for models, and explaining what factors are particularly important to a model
- technology, AutoML is likely to hugely expand the use of machine learning by organizations and take the creation of statistical models into a post-algorithmic age

Key Enterprise AI trends

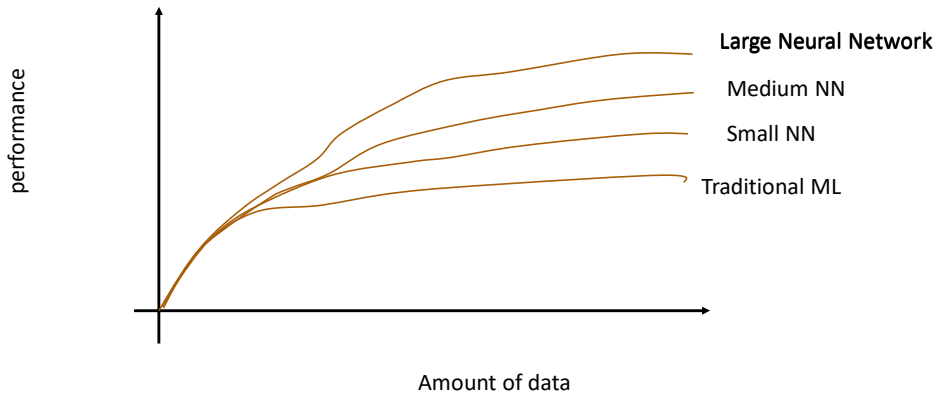
creation of AI centers of excellence and other management structures

- Organization start with roles like chief data and analytics officer, which then becomes responsible for AI, is well established in a plethora of companies including General Motors, JPMorgan Chase, Travelers, Wells Fargo, MetLife, Partners Healthcare, Marsh & McLennan, Walmart, CVS health etc.

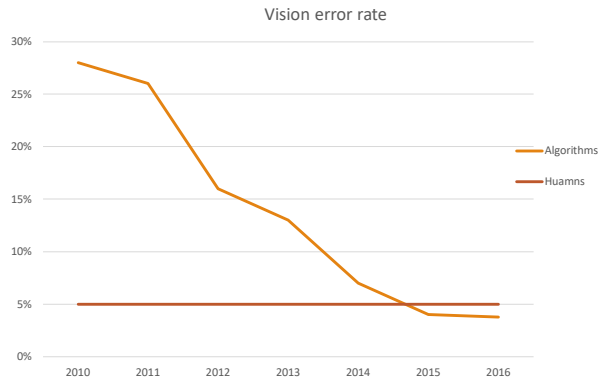
sparse data technologies

- One of the key requirements for many AI systems, particularly deep learning models, is voluminous data
- But many companies can't assemble enough data, and labeling it is almost always labor-intensive
- AI vendors are therefore working on a variety of technologies that make it possible to build high quality models and systems with smaller amounts of real world data
- Some of these systems, such as generative adversarial networks, create their own synthetic data

From traditional ML to Deep Learning



How performance has improved through ML / DL over time? E.g. Image Recognition



Machines have made real strides in distinguishing among similar looking categories of images

Source: Electronic Frontier Foundation

Supervised Learning

Case Udacity: Sebastian Thrun Co founder of Udacity noticed that some of the salespeople were much more effective than others when replying to inbound queries in a chat room

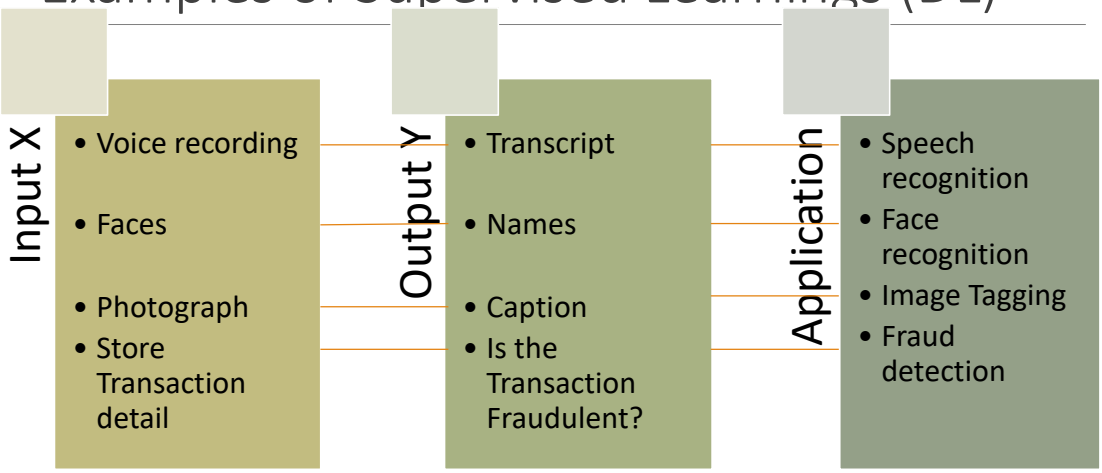
They carried out what is known as supervised learning. Interactions leading to sales were labeled as successes and all others were labeled failures

The data was used to predict what answers successful people were likely to give in response to certain very common inquiries and then shared those predictions with the other salespeople to nudge them toward better performance

After 1000 training cycles, the salespeople had increased their effectiveness by 54% and were able to service twice as many customers at a time

Udacity did not build a BOT but advised sales people how to improve their performance, could they have built a bot?

Examples of Supervised Learnings (DL)



Unsupervised learning

Alexa, Amazon's intelligent assistant that's in well over 100 million devices and works with over 60,000 appliances from 7,400 brands, gains new skills from contributions by the thousands of employees tinkering away at its backend systems. But there's a limit to what they can accomplish, owing to the way machine learning algorithms — the statistical models underpinning Alexa's decision-making — improve.

That's why scientists at Amazon's Alexa AI research division are pursing semi-supervised and unsupervised techniques, in which AI systems learn to make predictions without ingesting gobs of annotated data

Semi-supervised and unsupervised learning have their limitations, too, but both promise to supercharge Alexa's capabilities by imbuing a human-like capacity for inference.

Reinforced learning

In reinforced learning system the programmer specifies the current state of the system and the goal, lists allowable actions and describe the elements of the environment that constrain the outcomes for each of those action

For instance Microsoft uses reinforcement learning to select headlines for MSN.com news stories by “rewarding” the system with higher score when more visitors clicked on a link. The system tried to maximize its score on the basis of the rules its designers gave it. Of course reinforced learning optimizes the goal you explicitly reward. Specifying a goal correctly is critical

Deep Learning

Facebook AI tool Deep Text: to figure out meaning of words contextually from their conversation and then direct users toward products they might be interested in—one way of targeted Marketing(generating ads relevant to the individual’s likes.

Facebook AI tool DeepFace: uses facial recognition to identify people in photos (of course this tool is raising privacy concerns)

Netflix uses AI to generate recommendations for users based on viewing history

Amazon suggests other products that a consumer might be interested in based on purchase and search history.

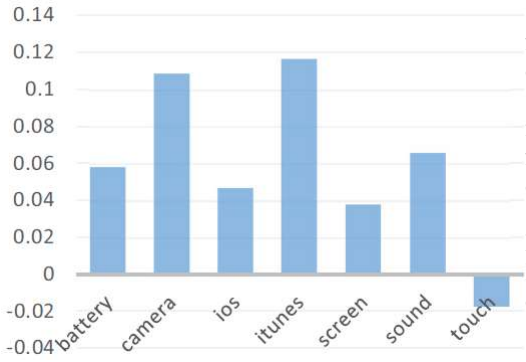
From Data to Action Leverage AI

1. Analyzing the current situation : By listening to customers sentiments , analyzing online conversation, rather than strictly focusing on own brand.

•Over 60% of people thought iPhone 6 as a "Good" mobile device
•The sentiment is in line with tech reviewers opinions
•"Excellent" sentiments are not obtained - as is usually the case since users tend to use simple descriptive terms about a device

Business Use cases:
Identifying a wider trend which can be used to either tailor products or marketing

Another sentiment analysis



•Camera and iTunes came out to be the two most important features - probably due to the DSLR level quality and user friendly features of iTunes
•Sentiments about touch and screen was on the lower or negative side which was identified to be due to a "bend" issue despite Apple suggesting it made phones more durable

Business Use cases:
Identifying customer delights, pain points and gaps in communication or satisfying customer expectations

Understanding Market and customers

Neurodata labs and robotics manufacturer unveiled a multimodal emotion detection for customer experience management. This AI system is able to analyze a combination of human activities (e.g. facial expressions, body gestures, voice, eye movement, and heart rate and determine consumer's emotional state.

This technology is tested by Russian Bank Rosbank at their call center, data collected like number of pauses in speech, changes in voice volume, total conversation time and converted in real time to customer satisfaction metric

Segmenting , targeting and positioning

Harley Davidson teamed up with US marketing firm Adgorithms to use its platform , Albert. Albert uses AI and ML automate marketing planning

HD provided Albert information on past customers, enabling the machine to create list of similar audience and expose the ad the customers resembling current buyers.

Platform helped in 40% increase in sales and ninefold increase in inbound calls