Integrate data analytics into business

Dr. Xiaomin Yang

1   Executive summary

Advanced data analytics has become a powerful catalyst of innovative solutions in industries, governments, and education sectors as may well-known applications such as e-commerce and media platforms demonstrated game-changing impacts. Recent failure of algorithm-powered instant house-buying business, however, reminds both business executives and data science educators of the challenges of translating algorithms into realizable business opportunities. This article analyzes lessons learned from the iBuying business cases and recommends a business-driven computer-human collaboration model for Integrating data-analytics solutions into enterprise value chain.

2   Data applications in business decisions

Applications of data analytics in businesses rapidly grow partially because of the reduced cost of data collection and processing, open accessibility to cloud-based analytical tools, as well as deep market penetration of online platforms.

A traditional area of data-driven applications is compiling a dashboard for informed decision-making. Business analysts collect, extract, and visualize different metrics and supporting data in a user-friendly format to provide decision makers an overview of reliable information relevant to their objectives. An insightful dashboard should be thoughtfully designed to report progress, trend, and prediction of key business performance indicators. Otherwise, presenting excessive data weakly aligned with strategic objectives will distract businesses from making timely decisions because humans tend to be indecisive when overwhelmed with unclear patterns and trends.

In recent years, advanced data analytics played a decisive role in many technology sectors, such as optimized marketing in an e-commerce business, customized video recommendations in the

media industry, and agile pricing for travel and hospitality booking websites. The application examples clearly demonstrated the power of data analytics for situations involving many stakeholders or customers. The success of those well-known applications is consistent with the nature of statistical analysis: finding correlations between variables to derive patterns and trends. Businesses use the insights derived from the algorithms to generate customized services for customers with similar preferences or needs.

Data analytics has become a powerful catalyst of innovative solutions in industries, governments, and education sectors as its applications appear endless, and its impacts are game-changing.

3   iBuying business case

Zillow's failure of algorithm-powered $6 billion instant house-buying business in 2022, however, reminds both business executives and data science educators of the challenges of translating data algorithms into realizable business opportunities.

Zillow Group is a tech real-estate advertising company headquartered in Seattle, Washington. Its business model is generating revenue by selling advertising on its website. Zillow website provides price estimates of these properties, in addition to basic real estate information. The popular price estimation service is one of the major features attracting users to visit Zillow.com. According to the company's website, A Zestimate is Zillow's estimated market value for a home, computed using a proprietary formula including public and user-submitted data, such as details about a home (bedrooms, bathrooms, home age, etc.), location, property tax assessment information and sales histories of the subject home as well as other homes that have recently sold in the area. Zillow has data on roughly 110 million homes across the United States.

Zillow claims, "the Zestimate's accuracy depends on location and the availability of data in an area. The Zestimate's median error rate for on-market homes nationwide is 3.2%, meaning Zestimate prices for half of all on-market homes are within 3.2% of the ultimate sale price, and half are not. For off-market homes, the median error rate is 7.52%."

The company emphasizes that a Zestimate is not an official appraisal but is a starting point to estimate a home's value.

In 2019, Zillow entered the digital home-buying business by launching a new service, Zillow Offers which made all-cash offers and purchased properties directly from home sellers. The service appeared to be rational and simple. Sellers only needed to answer some questions online, as well as took and uploaded a few pictures. Zillow would use the Zestimate computer algorithms to predict the fair market value of an eligible home and offer the seller a near-instant cash price that would include the cost of repairs and a small service fee (no more than 2% of the offering price to encourage future sellers). The company also deployed more than 100 pricing analysts to double-check the computer-generated numbers to reduce the risk of overpaying. If the seller accepted the offer, Zillow would buy the home, make a minor renovation, and resell it quickly on the market.

The algorithm-driven house-flipping business model, called iBuying, is based on the confidence that data science, fed by mega-data, can predict the specific price of any eligible house with pinpoint accuracy. Zillow held years of experience with its Zestimate algorithms and therefore, should be technically confident to incorporate relevant statistics into the Zillow Offers pricing model, which also includes photographic analysis. In addition, the company deployed many experienced human analysts to verify computation outcomes. Zillow Offers diligently took comprehensive data science best practices to maximize its chance of success.

According to a Wall Street Journal article, in 2021, Zillow Offers business grew 100% quarterly in terms of home purchase volume. Zillow bought 3,800 homes in the second, more than double the previous quarter. In the third quarter, it bought 9,680 homes.

On November 2nd, 2021, Zillow suddenly announced that the Zillow Offers business would permanently shut down. Zillow Group lost more than $500 million in 2021 from its Zillow Offers business. The company had to divest 7,000 homes with $2.8 billion value on its balance sheet after Zillow Offers ceased operations, according to Bloomberg.

The company attributed its Zillow Offers disaster to technical and market factors.

Zillow Group cofounder and CEO Rich Barton acknowledged that Zillow's pricing algorithm had failed to accurately predict fluctuations in home prices. In the company's [fiscal third-quarter earnings release](#) Mr. Barton said, "We've determined the unpredictability in forecasting home prices far exceeds what we anticipated, and continuing to scale Zillow Offers would result in too many earnings and balance sheet volatility." The company disclosed that Zillow Offers expected it would have to sell its homes at a 5% to 7% loss in the Fall of 2021.

Zillow struggled to find workers and building materials to repair the purchased properties amid a labor and supply shortage during the Covid-19 pandemic. The longer turnover period increased operation and financial costs, which had not been adequately addressed by the pricing algorithms.

Zillow pricing analysts complained that the company had retooled the algorithms to raise the analysts' suggested prices based on their "manual" analysis. Analysts whose job it was to confirm the prices of homes found that they were routinely overruled by the system.

Zillow Offers is not the house flipping practitioner experiencing a big loss from betting its future on an algorithm-based business model. Its iBuying competitor, Opendoor, lost $1.35 billion in 2022. Opendoor uses artificial intelligence and other technology to help it price and buy homes to quickly resell them at a profit margin. The company ended 2022 with an unsold inventory of 12,788 homes. It was expected to lose about 4% of a home on average for the properties sold in the 4th quarter of 2022.

4   What we could learn from the iBuying case

Zillow Group and Opendoor were confident, competent, and prepared to launch their iBuying businesses, respectively. The online real estate service providers had spent ten years polishing their fundamental home pricing algorithms used by millions of customers. Zillow Group collected property data on roughly 110 million homes across the United States. To fine-tune customized offers, iBuying services incorporated new graphic technologies into their pricing

methods. Zillow Offers even deployed 100 pricing analysts to double-check the computer-generated price numbers.

Zillow management and staff's reasoning about the failure appeared to be logical. To summarize key causes: 1) the short-term fluctuations of home prices are highly unpredictable; 2) rapid changes in labor and materials supply chain are not adequately addressed by the pricing method; 3) the Company overruled analysts by retooling the algorithms.

From a technical perspective, real estate experts explain that the complexity of the housing market makes it difficult to predict home prices months in advance. Data science faces difficulties in accurately incorporating some time-sensitive factors impacting a home's value into pricing algorithms.

Data-driven systems "trained" with historical data can predict future prices based on repeatable patterns and correlations captured by computing powers. To address unusual fluctuations and unprecedented market dynamics, computation analytics need real-time data feeds, generated dependent on the system, to continuously update algorithms.

It's reasonable to assume that those challenges are eventually resolvable as data science advances and computing power develops. However, many businesses proved that a thoughtfully developed business model can be successful before its computer-driven system becomes perfect.

4.1   Real estate business practices couldn't provide a real-time consumer & market feedback data to timely improve pricing algorithms

First, we can briefly compare how iBuying companies with e-commerce or online streaming companies use data in their business offers. For e-commerce, algorithms generate a customized shortlist of goods for a shopper to choose from. Buyers' purchase data is consistently processed to polish the algorithms to make more accurate recommendations next time, and more importantly, to help the company optimize its supply chain and inventory to minimize operation costs.

Online video services use algorithms to select customized media titles for an individual subscriber. Video streaming companies make their contents purchase decisions based on real-time viewers' favorite trends. Sophisticated media companies, such as Netflix and Disney, even incorporated detailed viewers preference data into their content creation decisions.

According to realtor.com, however, it takes an all-cash buyer 15 days to close a home deal and additional 30 days for sellers to move out. The renovation and reselling process will take more than two months. iBuying service providers must wait at least 100 days to collect a set of data for a home (purchase cost, renovation cost, and flipping revenue of the same property) to improve their algorithms through supervised learning. When the home prices increased rapidly in 2021, Zillow had to retool the system by embedding automatic price add-ons, including one called the "gross pricing overlay" that could add as much as 7%. It's not a good idea to code business judgment into computer systems.

Learning: Data models hold limited capabilities of using historical data to accurately predict future numbers under rapid changing business environment.

Learning: A business shall build transparent decision-making processes that clearly distinguish automated outcomes and business adjustment based on human judgement.

4.2 Zillow Offers business model inadequately integrates human cognitive behaviors with its pricing algorithms

E-commerce or streaming business practices enable companies and customers to make win-win decisions. E-commerce or streaming computation models generate multiple options for a consumer to decide. If a user could find and purchase products or content of interest, the online platform would meet its business goal.

Zillow Offers algorithms, however, provide only one purchasing price to a potential home seller, who must decide whether to accept the offer or not. Unfortunately, few sellers would make a home-sale decision without doing their research about the housing market. The traditional real-estate practices serve as an education process for a home seller to understand the comparable

housing market and eventually accept the property value. Most humans are risk-avert, which means we tend not to make decisions in uncertain situations. So only the home sellers who feel confident that Zillow Offers prices are higher than what they could get otherwise would decide to sell their properties to Zillow.

The reality proved this, Zillow claimed that only 10% of people who asked for a Zillow offer and eventually sold their home ended up selling it to Zillow. On average, Zillow overpaid 7% for these homes.

Learning: A business shall carefully consider customers' cognitive behaviors in its data-driven business model. Customers are human beings who are risk-avert in general and make intuitive decisions, different from what risk-neutral algorithms expect.

4.3   iBuying businesses could not efficiently manage the entire home flipping value chain

The entire value chain of the home flipping business includes several key components, such as purchasing a home at an agreeable price, renovating the property, and selling the home to a new buyer. According to realtor.com, selling a home involves about ten steps. Zillow could streamline its operations by cutting some duplicated steps, but flipping homes still involves a lengthy value chain.

iBuying business model is built around proprietary home-price prediction algorithms, which are the front end of the value chain. However, Zillow was less capable of efficiently managing some other elements of the value chain. For instance, Zillow claimed it struggled to find contractors and renovation materials amid a broader labor and supply shortage, leading to a longer home-hold period and therefore adding insurance and financing costs.

Learning: A business shall ensure capabilities in place to efficiently manage each key element of the value chain to realize the potential value of its computer-driven business model.

5   Framework for integrating data-analytics into a business

Computation-driven business models have disrupted many industries. Enterprises need to explore new opportunities to incorporate data analytics into their businesses to drive growth. But experts estimate failure rates that vary from 60% to 85%, according to a [Fortune article.](#)

A business-driven approach will provide the best chance of success to develop and integrate data analytics solutions into a business.

5.1   Be clear about business strategy and objectives

A business shall clearly define business strategy and objectives before considering data science.

[Business strategy](#) defines a company's specific approach to competing in an industry and clarifies the competitive advantages on which it will be based. The Five Forces framework, developed by Harvard University Professor Michael Porter, describes a method of analyzing the competitive environment and profitability of an industry to help a business understand which "forces" drive the distribution of economic value among the industry players, for example suppliers and customers. Porter also introduced four generic business strategies: cost leadership, differentiation, focused cost leadership, and focused differentiation. Cost leadership is establishing a competitive advantage by having the lowest cost of operation in the industry. Differentiation is the strategy that emphasizes offering a distinguishable, superior product than what competitors are providing. Focus is the strategy concentrating on a small segment of a defined market.

 Porter's Five Forces is considered [a static model](#) as it provides a snapshot of an industry' profitability drivers at a time point.

[Dynamic capability framework](#), developed by UC Berkeley Hass Business School Professor, David Teese, focuses on a firm's ability to integrate, build, and reconfigure internal and external competencies to address rapidly changing environments. Dynamic capabilities emphasize corporate agility, which is driven by its capacity to sense market opportunities and threats, take

opportunities, and develop competitive advantages by adjusting the enterprise's intangible and tangible assets.

In a situation where the outcome of a business strategy depends on not only its actions but also the actions of others such as competitors, game theory framework can help businesses analyze interactive situations among competing players. For example, game theory can be used to analyze real-world scenarios of price competition and product releases to predict their outcomes.

Business objectives are results that the business wants to deliver within a timeframe to achieve its strategy. A SMART goal method is usually used by a business to articulate Specific, Measurable, Achievable, Realistic, and Time-bound objectives. Business objectives could include financial performance, market penetration, operational excellence, as well as organization and capability growth.

5.2   Understand value chain and opportunities for improvement

Value chain refers to a progression of business activities and processes needed to create a valuable product or a service, and deliver it to a user with ensured satisfaction, which leads to improved customer loyalty and business profitability.

THE VALUE CHAIN

Harvard Business School Online
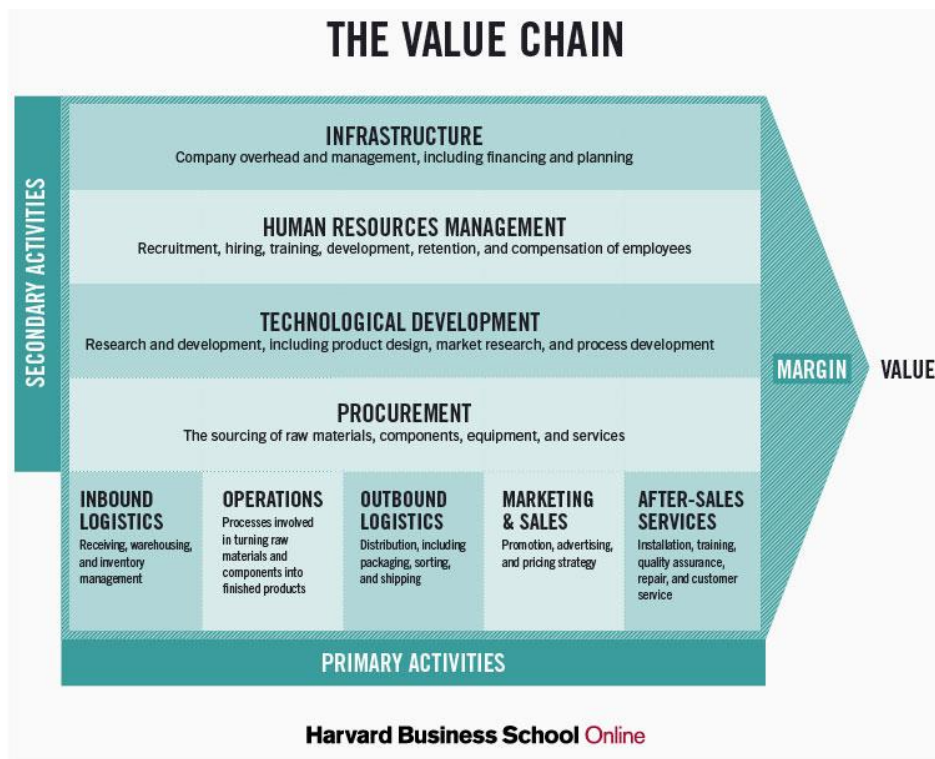
According to Porter, all these activities can be divided into two categories. Primary activities are those that go directly into the creation of a product or the performance of a service, including inbound logistics, operations, outbound logistics, marketing and sales, as well as after-sales services.

Secondary activities help primary activities become more efficient—effectively creating a competitive advantage—and are broken down into procurement, technological development, human resources management, infrastructure.

Value chain analysis is a method of evaluating activities in a company's value chain to understand opportunities for improvement at each step.

5.3 Establish a business-driven computer-human collaboration to develop and Integrate data-analytics solutions into enterprise value chain

How to translate data science into business challenges is a challenging topic for both business practitioners and educators. Integrating data science into business involves multidisciplinary

skills and mindsets, and therefore, requires close collaboration among all stakeholders to thoughtfully sign, rapidly develop and seamlessly incorporate data-driven solutions into business.

A computer-human collaborative team to develop and Integrate data-analytics solutions into enterprise value chain. The team includes subjective matter experts from corporate strategy and business unit, IT and data science, as well as sales and customer services, including other relevant functions.
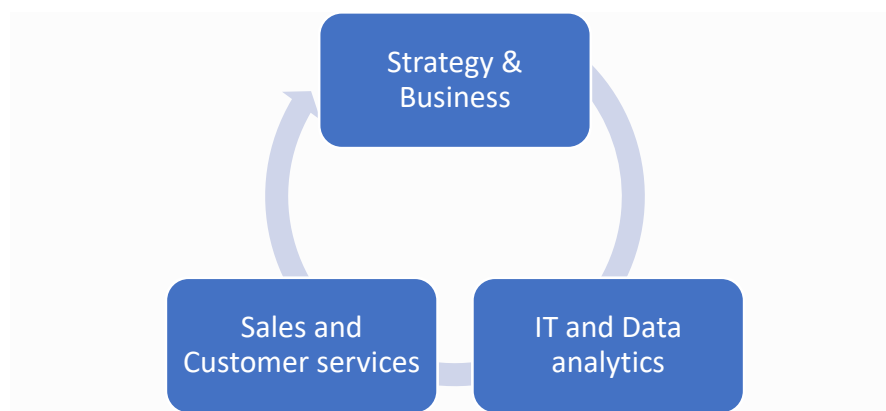
Computer-human collaboration

Corporate strategy management: high performing companies dedicate a strategy management function to ensure the effective execution of corporate strategy. The responsibilities of the strategic management function include ensuring that corporate roadmaps are translated into the business plans of all units, executing strategic initiatives to deliver long-term value, aligning employees' capability development plans, and their career goals and motivations, with strategic objectives, as well as rapidly adjusting enterprise's intangible and tangible assets to address dynamic competition. The function plays a coordination role to ensure computer-driven business initiatives are aligned with the company's strategy execution plan.

Business unit: a corporate assigns specific a product line to an independent organizational entity, a business unit which is a profit center responsible for its financial performance. A business unit typically implements discrete business processes to plan for market growth, new

products, and marketing to succeed in a unique competition environment. A business unit-level value chain analysis, in conjunction with [competition analysis](#), allows the organization to understand and define specific high-impact opportunities where data science can materialize to improve business profitability.

IT and Data analytics: industries have invested heavily in data science, artificial intelligence, and machine learning technologies and processes, but few companies successfully realized the full potential of data analytics. Based on survey studies, academic researchers and industrial experts recommended some practices to improve the chance of success in [cultivating a data culture](#), [implementing flexible data governance,](#) involving data subject matter experts in the [deployment of data solutions](#), etc. Data is a kind-of operating asset, and an organization should put formal governance in place to develop data assets with quality and integrity to enable every business unit, team, or individual to incorporate data-driven approaches into their decision-making processes, and eventually help the business make a profit by generating new revenue or reducing cost. Just like other new technologies, a data analytics solution needs to be justified with a validated business case, and its development-to-deployment lifecycle should follow appropriate governance processes to give the initiative a better chance of success and eventually realize its potential value.

Computer-human collaboration: a data-driven solution can create desired value for an enterprise only if the final service is successfully deployed with satisfying outcomes for both customers and the service provider. An enterprise should mobilize deployment and customer-facing SMEs to integrate the automation core with complementary features, particularly machine-human cooperation functions to create a superior product or service package which can provide excellent value to customers while meeting the enterprise's business ambitions. The users of data-driven services, e.g., iBuying, healthcare, and agile pricing, are human beings who think and behave differently from a machine.

[A hybrid computer–human workflow and team](#) can obviously combine the computing power and human intuitive capabilities to deliver successful businesses.

[After discussions, add a conclusion paragraph to wrap up]