# AI in Product Management

**Prof. Zillur Rahman**
**Department of Management Studies**
**Indian Institute of Technology, Roorkee**

## Lecture- 10
## Predictive Analytics in Market Research

Welcome to this NPTEL online certification course on Artificial Intelligence in Product Management. Now, we will talk about module 10, which is predictive analytics in market research. So, we are talking about this module, which is the last module. In this part 2, we have been talking about marketing research using AI. So, in this module, we will talk about how we can use predictive analytics in market research.

So, this is module 10, and with this, we will finish part 2. To have an overview of this module, in this module, we will understand how predictive analytics and different types of predictive models work. We will understand the role of predictive analytics in market research and the steps to implement predictive analytics. Then we will discuss various tools that can be used to perform predictive analysis and industry use cases of predictive analytics.

And then we will understand the challenges associated with predictive analytics. To give an introduction to predictive analytics, this is an area of statistics that deals with extracting information from data and using it to predict trends and behavior patterns. The basic premise is predicting what events will happen in the future. Businesses can use predictive analytics for a variety of use cases, such as forecasting consumer behavior, optimization of their marketing strategy, and operational efficiency. Predictive analysis involves a combination of data collection, data mining, statistical analysis, and predictive modeling. So, traditional analytics is mainly descriptive and diagnostic.

Descriptive analysis answers the question of what happened through historical data patterns and trends. The next level, diagnostic analytics, takes one step further in terms of why something happened. This time, using the data to uncover the root causes of these trends or patterns. Contrarily, predictive analytics is future-oriented. It answers what is expected to happen by building machines that predict future events based on models we have trained with past data.

Businesses that take a proactive approach are often able to anticipate and respond to future challenges and opportunities. Now, let us look at the types of predictive analytics models. Predictive analytic models are designed to assess historical data, discover patterns, observe trends, and use that information to predict future trends. Popular predictive analytic models include classification, clustering, and time series models. So, now let us look at the classification models.

Classification models fall under the branch of supervised machine learning models. These categories describe data based on historical data, describing relationships within a given data set. For example, this model can be used to classify customers or prospects into groups for segmentation purposes. Alternatively, it can also be used to answer questions with binary outputs, such as answering yes or no, or true and false. Popular use cases for this are fraud detection and credit risk evaluation.

Types of classification models use logistic regression, decision trees, random forests, neural networks, and Naive Bayes. The next is clustering models. Clustering models fall under unsupervised learning. They group data based on similar attributes. For example, an e-commerce site can use

the model to separate customers into similar groups based on common features and develop marketing strategies for each group. Common clustering algorithms include k-means clustering, mean-shift clustering, density-based spatial clustering of applications with noise, expectation-maximization clustering with Gaussian mixture models, and hierarchical clustering. Then come the time series models. Time series models use various data inputs at specific time frequencies such as daily, weekly, or monthly, etc.

It is common to plot the dependent variable over time to assess the data for seasonality trends, and they may indicate the need for a specific transformation and model types. Autoregressive moving average ARMA, ARIMA models are also frequently used time series models. As an example, a call center can use a time series model to forecast how many calls it will receive per hour at different times of the day. Now, we will look at predictive analytics in market research.

At its core, predictive analytics takes historical data and uses different types of algorithms to predict what could happen in the future with greater levels of accuracy. That way, businesses can predict changes in the market and what their customers want, and take steps based on the data so that they end up being more profitable. Being able to predict these trends allows companies to take action before the competition or react proactively to changes in their market. Deep insights into customer behavior.

Predictive analytics is one of the most significant factors. Using the data from historical interactions and purchase behaviors, companies can recognize different types of customers and align their marketing strategy according to the research. This way of targeting makes the experience much more customizable, resulting in higher rates of customer satisfaction and loyalty as well. Marketing strategies are fine-tuned with predictive analytics that uncover the best channels and messaging to approach consumers with. Boost marketing budgets and resources where they can deliver the most value.

This increased efficiency not only saves on cost but also allows marketing campaigns to be executed more efficiently. A more structured way. The benefits of predictive analytics in market research are substantial, as it helps businesses forecast future trends, behaviors, and outcomes based on historical data and advanced analytical techniques. Here is a detailed look at some of its key benefits. The first is accurate customer insight.

Predictive analytics enables businesses to gain deeper insights into customer preferences, behaviors, and by analyzing past customer interactions, purchases, and feedback. Marketers can predict future actions and trends. This helps in developing highly targeted marketing strategies that resonate with different customer segments. The second is optimized marketing campaigns. With predictive analytics, businesses can identify which marketing tactics work best for specific customer groups Marketers can run simulations or test scenarios before implementing full-scale campaigns, leading to better allocation of resources and maximizing ROI on marketing spends. By predicting outcomes, businesses can adjust campaigns in real-time to increase their effectiveness. The third is improved customer retention and loyalty.

Predictive analytics can identify early indicators of customer churn by recognizing patterns of disengagement or dissatisfaction. Businesses can take proactive measures to retain at-risk customers through targeted promotions or personalized offers, ultimately increasing customer lifetime value. The next is better sales forecasting. One of the most powerful applications of predictive analytics is in sales forecasting. By analyzing historical sales data and external factors like seasonality and economic trends, businesses can make more accurate predictions about future sales

volumes, helping in demand forecasting, inventory management, and setting achievable sales targets. Enhanced product development: Predictive analytics can inform product development by analyzing customer feedback, purchase patterns, and market trends. This helps businesses anticipate what features or products customers are likely to need in the future improving the relevance of a new product launch and minimizing the risk of product failure. Competitive advantage: By leveraging predictive analytics, businesses can anticipate market shifts and customer trends before competitors do. This allows them to stay ahead in the market, adjust their strategies dynamically, and seize new opportunities faster than competitors, giving them a significant edge.

Cost-efficient predictive analytics helps businesses minimize wasteful spending by identifying low-performing channels or strategies. By focusing efforts on high-performing customer segments and channels, companies can streamline their marketing budgets and reduce unnecessary costs. What are the steps to implement predictive analytics? Predictive analytics has the power to revolutionize your business, make accurate forecasts, and extract valuable business insights.

The necessary components are, first, to identify objectives, define clear goals for utilizing predictive analytics, and decide what you want to achieve with it, whether this is to improve sales predictions, increase customer retention, or optimize marketing campaigns. Then, collect and prepare data, gather the right resources with data, such as customer records, sales data, and market trends. Then, take this data, clean it up, and have it in the best place where you can analyze it. Next, use appropriate tools and techniques.

Select the appropriate predictive analytics tools based on your business needs. Popular tools that we have already talked about include IBM SPSS, SAS, and Microsoft Azure Machine Learning. Consider factors like ease of use, scalability, and integration with existing systems. Next, create and train predictive models.

Create models that predict the future from large datasets of the past. They make use of appropriate algorithms like linear regression or decision trees and train models to find patterns and predict answers. The fifth step is to test and validate models. Validate the accuracy of predictive models. Assess performance, use cross-validation, always evaluate and tune the model to get better predictions every time.

Now, how do we integrate predictive analytics into market research? Using predictive analytics in marketing research can be a tough task due to a series of challenges. There are data quality and access problems. Models are hungry beasts, and that means they need to be fed data to spit out predictions. We need to leverage these predictions only if the input data is valid.

At the same time, combining external data from various individual resources will typically be associated with certain discrepancies and inaccuracies. Additionally, a major hurdle to overcome is the sophistication of predictive models. However, building such models calls for expertise in data science and analytics, which might be missing in most market research teams. This skill gap can cause predictive analytics to be misused. Organizational inertia tends to push back on changes.

Many companies do not change their ways initially because the advances are not always obvious. In turn, their resistance can grind the predictive analysis implementation to a crawl and limit the ultimate impact on the business. To overcome these obstacles, businesses can adopt several practical solutions and best practices. Firstly, ensuring high-quality data is essential. This can be achieved.

By establishing robust data collection processes and using advanced tools for data cleaning and integration. Regular audits and updates to datasets can also help maintain accuracy and reliability. Another important solution is to invest in training and development. Companies can develop the necessary skill sets to create and manage predictive models either by upskilling current employees or by hiring specialized data scientists. Another way to gain this knowledge and support is by partnering with external experts or consultancy firms.

Now let us look at the tools to perform predictive analytics. The early days of analytics were dominated by analytics that helped enterprises understand what happened in the past. Descriptive analytics and diagnostic analytics. Developers commonly used various BI tools to create these models. Predictive analytics is a complementary field aimed at forecasting what could happen in the future by analyzing patterns and trends in past and current data.

Traditionally, predictive analytics was restricted to a small team of data analysts or data scientists. Predictive modeling was a complex process that required weeks or months of experimentation with different data sets, exploration of different hypotheses, and validation of different prototypes to find a model that showed value. This is now starting to change. With dramatic improvements in the capability of tools designed for both data analytics experts and regular business users. The term used to describe the various tools for building predictive models has also evolved over the years.

Today, they are commonly referred to as data science and machine learning tools. These tools are used to develop a variety of analytics and AI models used for descriptive, diagnostic, predictive, and prescriptive analytics. So, predictive analytics is just one aspect of these tools. In practice, users might not even directly refer to the term when applying predictive analytics to use cases. For example, a sales manager cares about a better lead scoring algorithm, a marketing manager wants a better click-through rate, and the finance team wants to reduce fraud.

In choosing the right predictive analytics tools for the job, it is essential to identify the enterprise, business, or function needs of your use cases. It might be the case that existing tools for business intelligence, analytics, or CRM already support your needs. It might also be worth investigating new data science or industry-specific tools that might have a better set of capabilities for your industry or business areas. Some predictive analytics focuses on more generic capabilities that can be applied across all industries, and other predictive analytics tools are specific to an industry or functional area.

Here are six of the most popular predictive analytic tools to consider. The first is Altair AI Studio. The Altair RapidMiner platform uses a comprehensive set of predictive analytic tools around its core data mining and text mining strengths. With Altair AI Studio supporting model development by data scientists, the platform's core capability simplifies extracting data from diverse sources. It cleans the data and incorporates it into various predictive modeling workflows.

Altair offers free trials of Altair AI Studio and the platform's other core products to help anyone get started and learn the basics. Our notebook feature simplifies the development of predictive analytics models for both novices and experts alike. The company also provides various augmented capabilities for data preparation with TurboPrep. It offers model generation with Auto Model and model operations for deployment, monitoring, and management of machine learning models. A generative AI extension enables users to build their own large language models and access open-source LLMs on the Hugging Face community.

The platform also supports various explainability and governance features when required. The second is H2O Driverless AI. A relative newcomer to predictive analytics, H2O has gained traction with a popular open-source offering. The company's H2O Driverless AI simplifies AI development and predictive analytics for both experts and citizen data scientists through open-source and custom recipes. Notable features include various automated and augmented capabilities for feature engineering, model selection, parameter tuning, natural language processing, and semantic analysis.

The company also offers a variety of capabilities to simplify the development of accessibility into predictive analytics models using causal graphs, local interpretable model-agnostic explanations, Shapley values, and decision-making sorobite methods. The third is IBM Watson Studio. IBM became a leading predictive analytics tools vendor with the acquisition of the Statistical Package for Social Sciences in 2009. SPSS was founded in 1975 and grew into one of the top statistical analysis packages over the years. IBM continued to innovate the vendor's core capabilities and integrated them into SPSS.

Modern Watson Studio on the IBM Cloud Pak for Data platform. This consolidated offering combines a broad range of descriptive, diagnostic, predictive, and prescriptive analytics functions. The platform simplifies predictive analytics for expert data scientists and improves collaborative data science for business users. The platform also includes various features to enhance responsible and explainable predictive models.

Next is Microsoft Azure Machine Learning. Microsoft has long been a leader in various analytical capabilities through its Power BI analytics platform in Excel, which can become the analytics front-end for the choice of most business users. The company's Azure Machine Learning complements these core tools with capabilities for managing the complete predictive analytics lifecycle. Supporting tools include Azure Data Catalog, Azure Data Factory, and Azure HDInsights.

The company supports all types of users, from expert data scientists to business subject matter experts. It also provides strong integration with its various application developments and RPA tooling, which makes it easier to deploy predictive analytics capabilities directly into applications and business workflows. The next is SAP's predictive analytics. SAP predictive analytics is a good example of how enterprise application platforms can extend their core offerings to support

predictive analytics workflows. The tool is a good choice for enterprises with extensive SAP deployments, particularly those looking to create predictive analytics for logistics, supply chain, and inventory management use cases. The current offerings were released in 2015 and built on two prior tools first released in 2012. The tool supports advanced and business users through various features that simplify data aggregation, predictive modeling, and model analysis across separate user interfaces. Then comes SAS.

SAS Institute is one of the oldest statistical analysis tool vendors. The first version of the company's tool started in 1966 as part of a US government initiative to improve data analysis for healthcare. The company was officially launched in 1972 after its government contract ran out. It has continued to innovate tools used by statisticians and data scientists.

The company is a clear leader in all kinds of analytic tools and techniques, including predictive analytics. More recently, the company has modernized its cold tool cells with various data science and machine learning workflows that take advantage of modern data stacks. Augmented workflow and simplified deployment. The company has hundreds of tools for various domains. Core offerings for predictive analytics include SAS Visual Data Science, SAS Data Science Programming, SAS Visual Data Decisioning, and SAS Visual Machine Learning.

The company also maintains strong relationships with leading cloud providers and enterprise software platforms to simplify predictive analytics development. And deployment across various workflows. Now, we will look at the predictive analytics industry users. Predictive analytics can be deployed across various industries for different business problems. Below are some industry use cases to illustrate how predictive analytics can inform decision-making within real-world situations. The first is banking. Financial services use machine learning and quantitative tools. To make predictions about their prospects and customers. With this information, banks can answer questions like who is likely to default on a loan, which consumers pose high or low risk, which customers are the most lucrative to target resources and marketing spends, and what spending is fraudulent in nature. The next is healthcare.

So, predictive analytics in healthcare is used to detect and manage the care of chronically ill patients, as well as track specific infections such as. Infections such as sepsis. Geisinger Healthcare used predictive analytics to mine health records to learn more about how sepsis is diagnosed and treated. Geisinger created a predictive model based on health records for more than 10,000 patients who have been diagnosed with sepsis in the past. The model yielded. Impressive results, correctly predicting patients with a high rate of survival.

Next, human resource teams use predictive analytics and employee survey metrics to match prospective job applicants, reduce employee turnover, and increase employee engagement. This combination of quantitative and qualitative data allows businesses to reduce their recruiting costs and increase employee satisfaction, which is particularly useful when labor markets are volatile. Next comes marketing and sales. While marketing and sales teams are very familiar with business intelligence reports to understand historical sales performance, predictive analytics enables companies to be more proactive in the ways they engage with their clients across the customer lifecycle.

For example, churn predictions can enable sales teams to identify dissatisfied clients sooner, enabling them to initiate conversations to promote retention. Marketing teams can leverage predictive data analysis for cross-sell strategies, and this commonly manifests itself through a recommendation engine on a brand's website. Then comes supply chain. Businesses commonly use predictive analytics to manage product inventories and set pricing strategies. This type of predictive analysis helps companies meet customer demand without overstocking warehouses.

It also enables companies to assess the cost of return on their products over time. If one part of a given product becomes more expensive to import, companies can project the long-term impact on revenues if they do or do not pass on additional costs to their customer base. Now, what are the challenges associated with predictive analytics? So, predictive analytics offers powerful insights, but implementing it effectively comes with several challenges. These challenges range from data-related issues to technical, operational, and organizational hurdles.

Below are some of the main challenges associated with predictive analytics. The first is data quality issues. Within this, the first issue is incomplete or inaccurate data. Predictive analytics relies on high-quality data. Poor data quality, such as missing, inconsistent, or inaccurate data, can lead to misleading predictions and unreliable models.

The second is data silos. In many organizations, data is spread across multiple systems and departments. When data is fragmented or stored in silos, it can be difficult to integrate and analyze holistically. This reduces the effectiveness of predictive models. The second issue is that of data privacy and security concerns.

The first concern here is regulatory compliance. Predictive analysis often involves the use of sensitive customer data, which raises privacy concerns. Regulations like GDPR, that is General Data Protection Regulation, and CCPA, that is California Consumer Privacy Act, impose strict rules on data collection and usage, requiring businesses to ensure that they handle personal data in a compliant manner. Data breaches and security risks. Storing and processing large amounts of customer and market data also increases the risk of data breaches.

Ensuring data security is critical but challenging, especially when working with external vendors or cloud-based systems. The third is model accuracy and reliability. Now, the first issue here is that of overfitting or underfitting. A major challenge in predictive analytics is ensuring that models are accurate and reliable. Overfitting occurs when a model is too complex and captures noise instead of relevant patterns.

While underfitting happens when a model is too simple to capture the complexities of the data. The second problem is also that of changing market conditions. Predictive models built on historical data may not always adapt well to changing market conditions, trends, or consumer behavior, making predictions less relevant over time. The next is that of complexity in implementation. The first issue here is the technical expertise required.

So, predictive analytics often requires specialized knowledge of statistics, data science, and machine learning. Organizations may struggle to find or train personnel with the required technical skills, making implementation difficult. The second problem here is that of complexity in model development. Developing accurate predictive models can be complex, requiring expertise in data processing, feature engineering, algorithm selection, and model tuning. This complexity increases with the scale of data and the sophistication of algorithms.

The fifth is integration with existing systems. So, the first problem here is that of legacy systems. Many organizations have legacy systems that may not be compatible with modern predictive analytics platforms. Integrating these systems with new tools or migrating data can be a time-consuming and expensive process.

Another problem that will arise here is that of data interoperability issues. Ensuring that data from different sources, that is, internal systems, third-party vendors, CRM platforms, etc., is interoperable and can be combined seamlessly for analysis is often a technical challenge. The next problem is that of biases in the predictive model. The first issue here is that of bias in data. If the historical data used to train predictive models contains demographic biases or selection biases, the resulting model will also be biased, leading to skewed predictions and unfair outcomes.

Another problem here is that of algorithmic biases. Even when the data is unbiased, certain algorithms may inadvertently introduce bias, especially if they rely on certain features or patterns that disproportionately affect specific groups. The next consideration is that of ethical considerations. The first issue here is the ethical use of data. Predictive analytics often raises ethical questions, particularly when it comes to using personal data without explicit consent.

Targeting vulnerable populations and making automated decisions that affect people's lives, for example, loan approvals and hiring. The next problem is that of fairness and discrimination. Predictive models must be carefully monitored to ensure they do not result in discriminatory practices. For example, in hiring or lending, biased models could unintentionally disadvantage certain demographic groups.

To conclude this module, we have discussed predictive analytics and different types of predictive analytic models. Then, we have understood the roles of predictive analytics in market research and the steps to implement predictive models. Then, we have discussed various tools that can be used to perform predictive analytics and industry use cases of predictive analytics. And finally, we have talked about the challenges associated with predictive analytics. These are some of the references from which the material for this module was taken.

Thank you.