AI in Product Management
Prof. Zillur Rahman
Department of Management Studies
Indian Institute of Technology, Roorkee

Lecture- 3
Role of AI in Product Management (Part 2)

Welcome to this NPTEL online certification course on artificial intelligence in product management. Now we are talking about Module 3, which is the role of AI in product management, Part 2. So we are discussing this part within the introduction to AI in product management. So this is Module 3 and Part 2 of Chapter 1. In this module, we will talk about how traditional organizations can digitally transform to support AI integration through the AI factory working model.

Then we will talk about what the various components of the AI factory working model are. Thereafter, we will see how product managers play a crucial role in this transformation from traditional to AI-assisted digital organizations and what the subfields of AI are. And their various types. So now let us start with the importance of AI in product management. In the ever-evolving tech industry, product management is experiencing a significant change due to the adoption of artificial intelligence.

Integrating AI into product management is no longer just a trend. It is a crucial strategy for achieving success. As companies transition into the AI-driven era, Updating software strategies is vital for excelling in product management. AI plays a key role in product-led organizations by fostering strategies that are centered on accuracy, personalization, and scalability.

The connection between AI and product-led growth, that is PLG, is clear. Presenting opportunities for strategic integration that can fuel innovation and offer a competitive edge. As AI's influence on modern product development grows, organizations must incorporate it into their strategies to remain competitive in today's fast-paced markets. Therefore, a product manager's usage of AI largely depends on how much or to what extent the organization has incorporated or integrated AI into its functioning.

The transformative impact of artificial intelligence is revolutionizing industries and fundamentally changing how organizations function. Although major players like Netflix have led the charge with their AI-driven operations, The creation of such capabilities is not restricted only to tech giants. Now let us look at AI-powered organizations, that is, the AI factory working model. The industrial revolution transformed manufacturing by introducing scalable and repeatable production methods.

Yet decision-making processes remained largely traditional and individualized. In today's AI-driven era, we are experiencing a profound shift where the industrialization of data collection, analytics, and decision-making is redefining the landscape. The AI factory represents a paradigm shift in decision-making and operational efficiency within the digital realm. It digitizes processes that were traditionally performed by employees and treats decision-making as an industrial process, leveraging the power of artificial intelligence.

By integrating AI, data processing, and analytics Organizations can leverage the power of scalable decision engines to drive innovation, improve customer experiences, and enhance overall performance. Now, we will look at the role of analytics in the AI factory. Analytics play a crucial role in the AI factory by systematically converting internal and external data into predictions, insights, and choices. These predictions guide

or automate various operational actions within the organization. By leveraging analytics, the AI factory enables superior scale, scope, and learning capacity within a digital firm. Next comes the digital operating models and AI factories. Digital operating models encompass the management of information flows, or the guidance of building, delivery, and operations of physical products. In both cases, the AI factory sits at the core, guiding critical processes and operational decisions.

Now, this shift allows humans to move to the periphery, focusing on tasks that require creativity, intuition, and complex problem-solving. Then comes the virtuous cycle of the AI factory. The Amazon flywheel cycle has growth at its core, but it is fueled by focusing on customers rather than profits. By providing a great customer experience at reasonable prices, it leads to happy customers who are more likely to come back. Thus, pushing the cycle, the Amazon Flywheel goes in two cycles that merge.

The first is the initial cycle, where the availability of third-party sellers leads to competition and a wider product selection. The second is that competition and a wider selection of products allow for lower prices, which fuels a positive customer experience in return. Positive customer experiences result in repeated purchases and That will bring in more traffic. Higher traffic brings in more third-party sellers and overall growth.

More sellers mean more selection choices. That leads to a better customer experience. More customers come, and therefore more sellers come. So, this is how this cycle works. The second cycle.

The company's growth opens the door for a low-cost structure. The lower cost structure allows for lower prices of products. Lower prices once again improve customer experience, and so on and so forth. So, that leads to a lower cost structure, which leads to lower prices, and then it again leads to customer experience. Now, what are the components of this AI factory?

The AI factory approach involves integrating analytics and AI into the core processes of the company, ensuring efficient and scalable operations. The essential components of an AI factory include the reverse data pipeline, algorithm development, experimentation platform, and software infrastructure. Now, what is the data pipeline? In AI factories, the data pipeline plays a crucial role in gathering, integrating, processing, and safeguarding vast amounts of user data. It ensures the smooth flow of data through the system.

Supporting real-time analytics and decision-making processes. Next is algorithm development, which focuses on generating predictions that drive critical operating activities. These algorithms continuously learn and adapt, refining recommendations and enhancing the user experience over time. Next is the experimentation platform.

The experimentation platform allows for rigorous testing of hypotheses and evaluation of the impact of suggested changes. It ensures that modifications have the intended effect and facilitates continuous improvement. The fourth is software infrastructure. The software infrastructure component of an AI factory provides a consistent and scalable framework for data delivery and connectivity.

It enables seamless interactions between internal and external users, ensuring efficient operations and enhancing the overall user experience. Now, you see that here we are gathering, cleaning, and normalizing the data. Then, this data goes to algorithm development, which is the second thing that includes supervised learning, unsupervised learning, reinforced learning, etc. Then, it goes to software infrastructure, which is software-enabled workflows, computing, storage, and analytics and then again, the cycle goes back. Here, we have this experimentation platform, and the product is deployed. Now, we will talk about the AI factory's role in data processing and analytics. The AI factory brings data processing and analytics, addressing the challenge of analyzing vast amounts of data.

By automating the process and integrating AI technologies, the AI factory forms the core of a digital operating model, driving decision-making, improving operational efficiency, and enabling continuous learning in modern digital forms. Now, what are the core values of the AI factory? Experimentation and improvement. The AI factory facilitates rigorous experimentation protocols to test hypotheses about customer behavior, competitive responses, and process variations. Data on usage, accuracy, and the impact of predictions are continuously fed back into the system for further processing and improvement.

This iterative process enables the AI factory to adapt and refine its algorithms over time. As the digital era continues to evolve, the AI factory will play a central role in shaping the future of businesses and unlocking new possibilities for success. Now, let us look at how product managers are at the forefront of the AI revolution. Product managers have become pivotal figures in the AI revolution. Uniquely positioned to leverage AI's potential to drive product development and spur innovation.

The traditional responsibilities of product managers overseeing product discovery, deployment, and delivery are being fundamentally reshaped by the integration of AI technologies. Rather than being passive observers, product managers are now active drivers and strategists. Guiding the incorporation of AI solutions into their products. As innovation leaders, they are tasked with seamlessly integrating AI into their product development processes, not only to keep up with technological progress but to set new standards for efficiency, functionality, and user experience. The potential of AI in the hands of product managers is immense.

Extending beyond process optimization to unlocking entirely new possibilities and approaches. From improving decision-making through machine learning to developing products that use natural language processing for more intuitive user interactions, AI empowers product managers to redefine what can be achieved. Now, this shift goes beyond technicalities. It is a fundamental change in the way product managers think. They must view AI both as a tool for developing

and shipping products and as a capability to enhance the value of their ecosystem. Now let us look at the various subfields of AI. AI's vast scope includes multiple subfields, each requiring a nuanced understanding to be effectively applied in product discovery and development. This is crucial for maximizing AI's potential and ensuring its optimal use in product management. So we will briefly discuss and understand the subfields of AI.

So one is natural language processing. At the intersection of AI and human communication lies natural language processing, a pivotal subfield that empowers machines to comprehend and interpret human language. NLP bridges the gap between the binary world of machines and the expressiveness of human communication. Tasks like language translation, Auto-correction and smart assistance on mobile devices epitomize the practical applications of NLP, making interactions with machines more intuitive and seamless.

Next comes visual perception. It involves a complex interplay of sensory inputs, cognitive processing, and contextual understanding to make sense of visual data. Perception within AI entails the analysis of data obtained from sensors, cameras, microphones, or similar input devices to comprehend the surrounding environment. This capability enables machines to simulate human perception by recognizing objects, comprehending speech, and interpreting visual and auditory signals. Next comes automatic

programming. Automatic programming, also known as code generation, refers to the process of automatically generating computer programs or source code using high-level specifications, algorithms, and machine learning techniques. This approach makes software development more efficient and less time-consuming, as it eliminates the need for human intervention in writing repetitive or complex code. Automatic programming tools can translate simplified inputs, such as user requirements or system models, into functional programs. Then comes knowledge representation.

Knowledge representation in AI refers to the way in which artificial intelligence systems store information. Organize and utilize knowledge to solve complex problems. It is a crucial aspect of AI, enabling machines to mimic human understanding and reasoning. Knowledge representation involves the creation of data structures and models that can efficiently capture information about the world, making it accessible and usable to AI algorithms for decision-making, inference, and learning. The next is intelligent robots.

Intelligent robots have a well-defined artificial brain which can arrange actions according to the purpose and also have sensors and effectors. Common technologies mainly include core components, robot-specific sensors, robot software, testing, safety, reliability, and other key common technologies. Key applications mainly include industrial robots, service robots, special environment service robots, and medical rehabilitation robots.

Demonstration applications are oriented toward industrial robots, medical rehabilitation robots, and other fields. Then comes automated reasoning. Automated reasoning is used to prove two things. First, they prove that a system design or implementation obeys its specifications. Second, they prove it works the way it was intended to.

Automatic reasoning does this by producing proofs in formal logic supported by mathematical theorems or known truths. Automatic reasoning uses mathematical logic-based algorithm verification methods to produce proofs of security. Or correctness for all possible behaviors. When you use automated reasoning, you first present the system with a problem statement. Then the automated reasoning system calculates and validates the assumption with the problem statement.

The software does this until it exhausts all options. Another subfield is machine learning. It is a subset of AI that relies on data and sophisticated algorithms, enabling machines to evolve and enhance their decision-making capabilities over time. A tangible manifestation of machine learning prowess is evident in the personalized product recommendations algorithmically crafted on e-commerce platforms.

Here, machine learning discerns integrated patterns from user behavior, optimizing the suggestions offered with each interaction. Machine learning algorithms can be categorized into three: supervised, unsupervised, and reinforcement learning. Supervised learning works by taking in clearly labeled data while being trained and using that to learn and grow. It uses the labeled data to predict outcomes for other data.

So, that is supervised learning. Unsupervised learning algorithms are given data that is not labeled. Unsupervised learning algorithms use that unlabeled data to create models and evaluate the relationship between different data points in order to give more insights into the data. Reinforcement learning algorithms learn by taking in feedback from the results of their actions. This is typically in the form of a reward.

A reinforcement algorithm is usually composed of two major parts: an agent that performs an action and the environment in which the action is performed. The cycle begins when the environment sends a state signal to the agent that cues the agent to perform a specific action within the environment. We will discuss some algorithm techniques of machine learning. One such technique is the decision tree. One of the most common supervised learning algorithms, decision trees get their name because of their tree-like structure, even though the tree is inverted.

The roots of the tree are the training databases, and they lead to specific nodes which denote a test attribute. Nodes often lead to other nodes. A node that does not lead onward is called a leaf. Decision trees classify all the data into decision nodes. They use a selection criterion called attribute selection measures, which takes into account various measures, for example, entropy, gain ratio, information gain, etc.

Using the root data and following the ASM, which is the attribute selection measure, the decision tree can classify the data. It is given by following the training data into sub-nodes until it reaches the conclusion. So, this is a decision node, this is the subtree to further decision nodes, this is a leaf, this is a leaf, this is a leaf, this is again a decision node. So, ultimately, all these decision nodes will end in a leaf node. So, here you are with friends, yes, it is windy, no cold, above par, no below par branch.

Now, these are leaf nodes: walk or cart, walk, above or cold. Another machine learning technique is the random forest. The random forest algorithm is actually a broad collection of different decision trees, leading to its name. The random forest builds different decision trees and connects them to gain more accurate results. This can be used for both classification and regression types of supervised learning.

While a solo decision tree has one outcome, and a narrow range of groups, the forest assures a more accurate result with a bigger number of groups and decisions. It has the added benefit of adding randomness to the model by finding the best feature among a random subset of features. Overall, these benefits create a model that has wide diversity, which many data scientists favor. So, as we can see from the diagram, the results of decision trees 1, 2, and 3 are combined, which are then averaged out.

Or the majority is considered as the final result. So, here we have this dataset; this is decision tree 1, results 1, 2, 3 averaging, and then we get the final result. Next comes the principal component analysis. As the number of features or dimensions in a dataset increases, the amount of data required to obtain a statistically significant result increases exponentially. Principal component analysis.

It works on the condition that while the data in a higher-dimensional space is mapped to data in a lower-dimensional space, the variance of data in the lower-dimensional space should be maximum. It is used to examine the interrelationship among a set of variables, and it is also known as a general factor analysis where regression determines the line of best fit. The main goal is to reduce the dimensionality of a dataset by finding a new set of variables that is smaller than the original set of variables, retaining most of the sample's information. and useful for the regression and classification of data.

So, this is PC2; this is the, these are area and radius PC1 principal component; this is the variance. So, the transformation goes from 2D to 1D; PC2, PC1 is greater than PC2. Next comes the k-means, which sorts the remaining data points into clusters based on their proximity to each other and the centroid data points for each cluster. The algorithm takes the unlabeled datasets as inputs, divides the dataset into k numbers of clusters, and repeats the process until it finds the best clusters. The value of k should be predetermined in this algorithm.

The k-means clustering algorithm mainly performs two tasks. One is to determine the best value of K center points or centroids by an iterative process. The second is to assign each data point to its closest K center. Those data points which are near to the particular K center create a cluster. K means that the working of a K-means algorithm is as follows.

First is to select the number of K's to decide the number of clusters. Second is to select random K points or centroids. Assign each data point to their closest centroid which will form the predefined K clusters. Then calculate the variance and place a new centroid in each cluster. Repeat the third step which means reassigning each data point to the new closest centroid of each cluster.

So, this is before the k-means. Now, these are the various clusters. Now, this is after k-means. So, now it becomes much clearer to understand the various clusters. The next comes the k-nearest neighbor.

The KNN algorithm can be used for regression as well as for classification, but mostly it is used for classification problems. Suppose there are two categories, category A and category B, and we have a new data point x1. So, in which of these categories will this data point lie? To solve this type of problem, we need the KNN algorithm. With the help of KNN, we can easily identify the category or class of a particular dataset.

So, this is before KNN: this is the new data point, category B, and category A. Now, after KNN, this is category B, and the new data point is assigned to category 1. And this is category A. Next comes linear regression. Linear regression is a supervised learning AI algorithm used for regression modeling. It is mostly used for rediscovering the relationship between data points, predictions, and forecasting. Much like support vector machines, it works by plotting pieces of data on a chart with the x-axis as the independent variable and the y-axis as the dependent variable.

The data points are then plotted out in a linear fashion. To determine their relationships and forecast possible future data. Linear regression is one of the easiest and most popular machine learning algorithms. It is a statistical method that is used for predictive analysis. Linear regression makes predictions for continuous, real, or numeric variables such as sales, salary, age, product price, etc.

The linear regression algorithm shows a linear relationship between a dependent variable, that is y, and one or more independent variables, that is x, hence called linear regression. So, this is the independent variable x, this is the dependent variable y, and these are the data points fitted on this regression line. Since linear regression shows a linear relationship, it finds how the value of the dependent variable changes according to the value of the independent variable. The linear regression model provides a sloped straight line

representing the relationship between the variables. Next comes logistic regression. The logistic regression algorithm usually uses a binary value, that is 0 or 1, to estimate values from a set of independent variables. The output of logistic regression is either 1 or 0, yes or no. An example of this would be a spam filter in email.

The filter uses logistic regression to mark whether an incoming email is spam, 0, or not, 1. Logistic regression is only useful when the dependent variable is categorical. That is, either yes or no. The logistic regression model is based on the logistic function, which is a type of S-shaped curve that maps any continuous input to a probability value between 0 and 1. The logistic function allows us to model the relationship between the independent variables and the probability of the dependent variable taking on the value of 1.

The logistic regression model estimates the coefficients of the independent variables that are most predictive. of the independent variable. These coefficients are used to create a linear equation that is then transformed by the logistic function to produce a probability value for the dependent variable. taking on the value of 1. So this is that S-shaped curve that we have talked about earlier.

Logistic regression is commonly used in fields such as healthcare, marketing, finance, and social sciences to predict the likelihood of an event occurring, such as whether a patient has a certain disease or whether a customer will buy a product. The next comes support vector machines. The support vector machine, that is the SVM algorithm, is another common AI algorithm that can be used for either classification or regression. but is most often used for classification. SVM works by plotting each piece of data on a chart in n-dimensional space where n is the number of features.

Thus, the algorithm classifies the data points by finding the hyperplane that separates each class. There can be more than one hyperplane. So, these are the positive hyperplane, maximum margin, negative hyperplane and these are the support vectors. The main objective of a support vector machine is to segregate the given datasets in the best possible way.

The distance between the nearest points is known as the margin. The objective is to select a hyperplane with the maximum possible margin between support vectors in the given dataset. SVM searches for the maximum marginal hyperplane. Generate hyperplanes that segregate the classes in the best way. The figure on top shows three hyperplanes: black, blue, and orange.

Here, the blue and orange have higher classification errors, but the black is separating the two classes correctly. Select the right hyperplane with the maximum segregation from the nearest data points, as shown in the figure at the bottom. Next comes neural networks. A neural network is a machine learning program or model that makes decisions in a manner similar to the human brain by using processes that mimic the way biological neurons work together to identify phenomena, weigh options, and arrive at conclusions.

Every neural network consists of layers of nodes or artificial neurons: an input layer, one or more hidden layers, and the output layer. So, this is the input layer, multiple hidden layers, and this is the output layer. Each node connects to others and has its own associated weight and threshold. If the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network. Otherwise, no data is passed along to the next layer of the network.

Neural networks rely on training data to learn and improve their accuracy over time. Once they are fine-tuned for accuracy, they are powerful tools in computer science and artificial intelligence, allowing us to classify and cluster data at a high velocity. Tasks in speech recognition or image recognition can take minutes versus hours. One of the best-known examples of a neural network is Google's search algorithm. We will discuss a few types of neural networks in the next slides.

The first of these neural networks is the Boltzmann machine. A Boltzmann machine is an unsupervised deep learning model in which every node is connected to every other node. It is a type of recurrent neural network, and the nodes make binary decisions with some level of bias. These machines are not deterministic deep learning models. They are stochastic or generative deep learning models.

There are two types of nodes in the Boltzmann machine: visible nodes, which are measured, and hidden nodes, which are not measured. Although the node types are different, the Boltzmann machine considers them the same, and everything works as one single system. So, these are the visible nodes, and then there are the hidden nodes. The training data is fed into the Boltzmann machine, and the weights of the system are adjusted accordingly. The training data is fed into the Boltzmann machine and the weights of the systems are adjusted accordingly.

Boltzmann machines help us understand abnormalities by learning about the working of the system in normal conditions. Next comes multiple layer perceptrons. The multiple layer perceptron, that is MLP, is a type of artificial neural network consisting of multiple layers of neurons. The neurons in the MLP typically use non-linear activation functions. Allowing the network to learn complex patterns in data.

MLPs are significant in machine learning because they can learn non-linear relationships in data, making them powerful models for tasks such as classification, regression, and pattern recognition. There is one neuron or node. It has one output layer with a single node for each output, and it can have any number of hidden layers. And each hidden layer can have any number of nodes.

If it has more than one hidden layer, it is called a deep artificial neural network. An MLP is a typical example of a feedforward artificial neural network. So, deep learning can process. Next is deep learning. Deep learning can process an extensive array of data types, including images and sounds, including image and sound.

Its application transcends the contextual domain. Finding resonance in groundbreaking advancements such as the development of driverless cars, showcasing its abilities in object retention and decision-making, ushering in a new era of sensory understanding for machines. The next slide will briefly explain how deep learning neural networks work. Next comes recurrent neural networks. It is a type of neural network where the output from the previous step is fed as input to the current step.

In traditional neural networks, all the inputs and outputs are independent of each other. Still, in cases where it is required to predict the next word of a sentence, the previous words are required, and hence there is a need to remember the previous words. Thus, RNN comes into existence, which solved this issue with the help of a hidden layer. So, this is what a recurrent neural network looks like. The main feature of RNN is its hidden state, which remembers some information about a sequence.

The state is also referred to as a memory state since it remembers the previous inputs into the network. It uses the same parameters for each input as it performs the same task on all inputs or hidden layers to produce the output. This reduces the complexity of parameters, unlike other neural networks. Next comes the convolutional neural network. The convolutional neural network, that is CNN, also known as ConvNet, is a specialized type of deep learning algorithm mainly designed for tasks that necessitate object recognition, including image classification, detection, and segmentation.

CNNs are employed in a variety of practical scenarios such as autonomous vehicles, security cameras, and others. CNNs mimic the human visual system but are simpler, lacking its complex feedback mechanism and relying on supervised learning rather than unsupervised learning, driving advances in computer vision despite these differences. Then comes Generative Adversarial Network. So, this GAN represents a cutting-edge approach to generative modeling within deep learning, often leveraging architectures like convolutional neural networks. The goal of generative modeling is to

autonomously identify patterns in input data, enabling the model to produce new examples that feasibly resemble the original dataset. Generative adversarial networks can be broken down into three parts. The first is generative. To learn a generative model, we describe how data is generated in terms of probabilistic models. Second is adversarial; the generative result is compared with the actual image in the dataset.

A mechanism known as a discriminator is used to apply a model that attempts to distinguish between real and fake images. And the third is networks; you use deep networks as artificial intelligence algorithms for training purposes. So, this is what it looks like; here we have a latent random variable generator. real data samples, discriminator, and condition. Is it correct?

The next term is the deep belief network. They are sophisticated artificial neural networks used in the field of deep learning. A subset of machine learning, they are designed to discover and learn patterns within large datasets automatically. We can imagine them as multi-layered networks where each layer is capable of making sense of the information received. From the previous one, gradually building up a complex understanding of the overall data.

It is useful for tasks like image and speech recognition where the input data is high-dimensional and requires a deep level of understanding. The architecture of DBN also makes them good at unsupervised learning, where the goal is to understand and label input data without explicit guidance. This characteristic is particularly useful in scenarios where labeled data is scarce or when The goal is to explore the structure of the data without any preconceived labels. Then another field, a subfield of AI, is computer vision.

Computer vision is a field of artificial intelligence that uses machine learning and neural networks to teach computers and systems to derive meaningful information from digital images, videos, and other visual inputs to make recommendations or take actions when they see defects or issues. If AI enables computers to think, computer vision enables them to see, observe, and understand. The next subfield is generative AI. Unleashing the power of creativity within AI, generative AI, exemplified by large language models like GPT-4 or Google's BARD, serves as an artistic force in the digital realm. This subfield thrives on generating content in response to prompts, showcasing versatility in creating anything that

Anything from concise report summaries to engaging promotional emails. Generative AI is an innovative tool that simplifies the creative capacities of machines. Now, in this figure, we summarize all these methods. This is the AI ecosystem. Here, we have generative AI and large language models.

So, to conclude this module, we have understood the digitization of traditional organizations to support AI ecosystems with the help of an AI factory working model and its various components. Thereafter, we have discussed how product managers play a pivotal role in this transformation journey from traditional to the incorporation of AI. Finally, we have understood the subfields of AI and their various types, which are all parts of the AI ecosystem. And these are some of the references from which the material for this module was taken. Thank you.