AI: Strategy + Marketing (MGT 853)

AI in Retail + Fashion (Session 9)

Vineet Kumar

Yale School of Management

Spring 2025

Course Logistics

1. Guest Speaker – Please sign up for lunch (need accurate count)

Course Logistics

1. Guest Speaker – Please sign up for lunch (need accurate count)
2. Project feedback ?

Course Logistics

1. Guest Speaker – Please sign up for lunch (need accurate count)
2. Project feedback ?
3. Project Slides – Please upload them by 9 am of your presentation date

Course Logistics

1. Guest Speaker – Please sign up for lunch (need accurate count)
2. Project feedback ?
3. Project Slides – Please upload them by 9 am of your presentation date
   - Everyone can update until May 1 at 9 am for grading

Course Logistics

1. Guest Speaker – Please sign up for lunch (need accurate count)
2. Project feedback ?
3. Project Slides – Please upload them by 9 am of your presentation date
    - Everyone can update until May 1 at 9 am for grading
4. All assignments are now complete

Building Data Science Products at the New York Times

Pablo Romano

Machine Learning Scientist
Ads Data Science

Date: Wed, April 30, 2025
Time: 11:45 am - 12:45 pm
Location: 4200 Qian and Yu Classroom
Evans Hall, 165 Whitney Ave

Pablo Romano is a Lead Machine Learning Scientist within the Advertising Mission at the New York Times. He supports the research, development, and integration of NLP systems and LLMs across several Advertising data products. He has been actively involved in the emerging application of Generative AI to enhance products across the Ads business.

He holds a PhD in Physical Chemistry from the University of Oregon where he developed novel ML methods to study large scale molecular simulations of DNA.

His presentation will cover how ML and AI products are scoped and developed at the New York Times, as well as cover some case studies for our most successful data products.

The New York Times
NEW YORK TIMES ADVERTISING • 2025 SLATE

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?
4. Map out the decisions to ML (prediction) problems

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?
4. Map out the decisions to ML (prediction) problems
5. What is the average unit sales of products (SKU-size) for each store per year?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?
4. Map out the decisions to ML (prediction) problems
5. What is the average unit sales of products (SKU-size) for each store per year?
6. What impact does the above have on the ML system?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?
4. Map out the decisions to ML (prediction) problems
5. What is the average unit sales of products (SKU-size) for each store per year?
6. What impact does the above have on the ML system?
7. How can **Human** $\Longleftrightarrow$ **AI** interaction create value for the firm?

Miroglio – Questions for Discussion

1. What are some of the unique challenges in the fashion retail industry?
2. What business problems and decisions is Miroglio faced with?
3. What are some non-ML approaches to solving these problems?
4. Map out the decisions to ML (prediction) problems
5. What is the average unit sales of products (SKU-size) for each store per year?
6. What impact does the above have on the ML system?
7. How can **Human** $\Longleftrightarrow$ **AI** interaction create value for the firm?
8. How can we operationalize the human to AI interface?

How many Units of Items per Store per Year?

AI model with Store manager input

In class Exercise

Small Group Discussion

Forms groups of 3 or so, and explore the following questions.

- How would you design an AI model that would include store manager input?

AI model with Store manager input

In class Exercise

Small Group Discussion

Forms groups of 3 or so, and explore the following questions.

- How would you design an AI model that would include store manager input?
- Detail the form of the manager input. Do you want X or Y variables? or both?

AI model with Store manager input

In class Exercise

Small Group Discussion

Forms groups of 3 or so, and explore the following questions.

- How would you design an AI model that would include store manager input?
- Detail the form of the manager input. Do you want X or Y variables? or both?
- How would you validate that the input is correct?

AI model with Store manager input

In class Exercise

Small Group Discussion

Forms groups of 3 or so, and explore the following questions.

- How would you design an AI model that would include store manager input?
- Detail the form of the manager input. Do you want X or Y variables? or both?
- How would you validate that the input is correct?
- How would you think about incentives for the manager to share this information?

Miroglio – What Happened

- After much internal debate, Francesco Cavarero and Davide Garelli decided to partner with Evo Pricing (Evo) to build an artificial intelligence (AI) system for demand forecasting and inventory replenishment for Miroglio's Elena Mirò brand.

Miroglio – What Happened

- After much internal debate, Francesco Cavarero and Davide Garelli decided to partner with Evo Pricing (Evo) to build an artificial intelligence (AI) system for demand forecasting and inventory replenishment for Miroglio's Elena Mirò brand.
- Evo CEO Fabrizio Fantini and his team of data scientists worked in close collaboration with Garelli's team and with Miroglio's Chief Data Officer, to come up with the following system.

Miroglio – Designing AI system - A

Initial Inventory Allocation Decisions:

1. First decision depended on the accuracy of and confidence in the new replenishment system. Based on the system described below, Evo decided to allocate 60% ($\downarrow$ from 85%) of the total inventory at the start of the season.

Miroglio – Designing AI system - A

Initial Inventory Allocation Decisions:

1. First decision depended on the accuracy of and confidence in the new replenishment system. Based on the system described below, Evo decided to allocate 60% (↓ from 85%) of the total inventory at the start of the season.
2. Second followed Miroglio's segmentation of store clusters, ensuring that its larger stores received more inventory of the same assortment.

Miroglio – Designing AI system - A

Initial Inventory Allocation Decisions:

1. First decision depended on the accuracy of and confidence in the new replenishment system. Based on the system described below, Evo decided to allocate 60% ($\downarrow$ from 85%) of the total inventory at the start of the season.
2. Second followed Miroglio's segmentation of store clusters, ensuring that its larger stores received more inventory of the same assortment.
3. Third decision of allocating specific items to a store was done by first identifying features of each item (e.g., its fabric, category, color, price point etc.). Item matched to previous items with similar features sold in store.

Miroglio – Designing AI system - A

Initial Inventory Allocation Decisions:

1. First decision depended on the accuracy of and confidence in the new replenishment system. Based on the system described below, Evo decided to allocate 60% (↓ from 85%) of the total inventory at the start of the season.
2. Second followed Miroglio's segmentation of store clusters, ensuring that its larger stores received more inventory of the same assortment.
3. Third decision of allocating specific items to a store was done by first identifying features of each item (e.g., its fabric, category, color, price point etc.). Item matched to previous items with similar features sold in store.
4. Finally, the model computed the initial allocation of the item as the weighted first four weeks' sales volume of those other items.

Miroglio – Designing AI system - A

Initial Inventory Allocation Decisions:

1. First decision depended on the accuracy of and confidence in the new replenishment system. Based on the system described below, Evo decided to allocate 60% (↓ from 85%) of the total inventory at the start of the season.
2. Second followed Miroglio's segmentation of store clusters, ensuring that its larger stores received more inventory of the same assortment.
3. Third decision of allocating specific items to a store was done by first identifying features of each item (e.g., its fabric, category, color, price point etc.). Item matched to previous items with similar features sold in store.
4. Finally, the model computed the initial allocation of the item as the weighted first four weeks' sales volume of those other items.
5. Expensive black blouse was more likely to be allocated to stores that sold many black items (not just blouses), had higher sales of blouses (of all colors), and sold more expensive items

Miroglio – Designing AI system - B

Inventory Replenishment Decisions:

1. Fantini and his team knew that individual Miroglio stores sold few units in each size of SKU $\Longrightarrow$ data at disaggregated level too sparse to predict accurately

Miroglio – Designing AI system - B

Inventory Replenishment Decisions:

1. Fantini and his team knew that individual Miroglio stores sold few units in each size of SKU $\Longrightarrow$ data at disaggregated level too sparse to predict accurately
2. Aggregated past sales of an item in a given category (e.g., women's pants) across all sizes to first build a forecast at an item-store level

Miroglio – Designing AI system - B

Inventory Replenishment Decisions:

1. Fantini and his team knew that individual Miroglio stores sold few units in each size of SKU $\Longrightarrow$ data at disaggregated level too sparse to predict accurately
2. Aggregated past sales of an item in a given category (e.g., women's pants) across all sizes to first build a forecast at an item-store level
3. Removed the effects of markdowns on past sales and estimated a seasonality coefficient. Then adjust item-store forecasts (How?)

Miroglio – Designing AI system - B

Inventory Replenishment Decisions:

1. Fantini and his team knew that individual Miroglio stores sold few units in each size of SKU $\Longrightarrow$ data at disaggregated level too sparse to predict accurately
2. Aggregated past sales of an item in a given category (e.g., women's pants) across all sizes to first build a forecast at an item-store level
3. Removed the effects of markdowns on past sales and estimated a seasonality coefficient. Then adjust item-store forecasts (How?)
4. Finally, they arrived at the size allocation for each item-store forecast by giving equal weight to two factors: relative frequency of a certain size for an item across all stores; and relative frequency of a size for a store across all items

Miroglio – Designing AI system - B

Inventory Replenishment Decisions:

1. Fantini and his team knew that individual Miroglio stores sold few units in each size of SKU $\Longrightarrow$ data at disaggregated level too sparse to predict accurately
2. Aggregated past sales of an item in a given category (e.g., women's pants) across all sizes to first build a forecast at an item-store level
3. Removed the effects of markdowns on past sales and estimated a seasonality coefficient. Then adjust item-store forecasts (How?)
4. Finally, they arrived at the size allocation for each item-store forecast by giving equal weight to two factors: relative frequency of a certain size for an item across all stores; and relative frequency of a size for a store across all items
5. Leveraging Image Data: Analyze images of fashion items to abstract key attributes (these might include not only attributes typically identified by humans, such as shape and design, but also many other unique aspects that might be hard for people to describe). It would then correlate these image attributes with sales.

Miroglio – Designing AI system - C

Black Box or Explainable? Fantini wondered whether to propose that Miroglio management adopt the simple, explainable, but less accurate approach, or the state-of-the-art AI method which provided far better forecasts but could be perceived as a "black box." Implementing a state-of-the-art approach could establish Evo as the AI and machine learning pioneer in the fashion industry, but would it create resistance from Miroglio's management?

Miroglio – Designing AI system - D

Including Manager Input on Customers and Local Trends:

1. Idea: providing each store manager with a **non-monetary budget** that they could use to accept or replace replenishment items that Miroglio's HQ proposed.

Miroglio – Designing AI system - D

Including Manager Input on Customers and Local Trends:

1. Idea: providing each store manager with a **non-monetary budget** that they could use to accept or replace replenishment items that Miroglio's HQ proposed.
2. Garelli explained: "Our Head of Buying has a budget that she uses to decide how many blue jeans or black blouses to buy based on a demand forecast. However, she uses her judgment to modify these quantities from the demand forecast. We can offer a similar budget to store managers and effectively make them the buying heads of their own stores."

Miroglio – Designing AI system - D

Including Manager Input on Customers and Local Trends:

1. Idea: providing each store manager with a **non-monetary budget** that they could use to accept or replace replenishment items that Miroglio's HQ proposed.
2. Garelli explained: "Our Head of Buying has a budget that she uses to decide how many blue jeans or black blouses to buy based on a demand forecast. However, she uses her judgment to modify these quantities from the demand forecast. We can offer a similar budget to store managers and effectively make them the buying heads of their own stores."
3. Head of Retail Carlo Tibaldi was very excited about this idea: approach would empower store managers and encourage them to embrace the analytical model.

Miroglio – Designing AI system - D

Including Manager Input on Customers and Local Trends:

1. Idea: providing each store manager with a **non-monetary budget** that they could use to accept or replace replenishment items that Miroglio's HQ proposed.
2. Garelli explained: "Our Head of Buying has a budget that she uses to decide how many blue jeans or black blouses to buy based on a demand forecast. However, she uses her judgment to modify these quantities from the demand forecast. We can offer a similar budget to store managers and effectively make them the buying heads of their own stores."
3. Head of Retail Carlo Tibaldi was very excited about this idea: approach would empower store managers and encourage them to embrace the analytical model.

Miroglio – Designing AI system - D

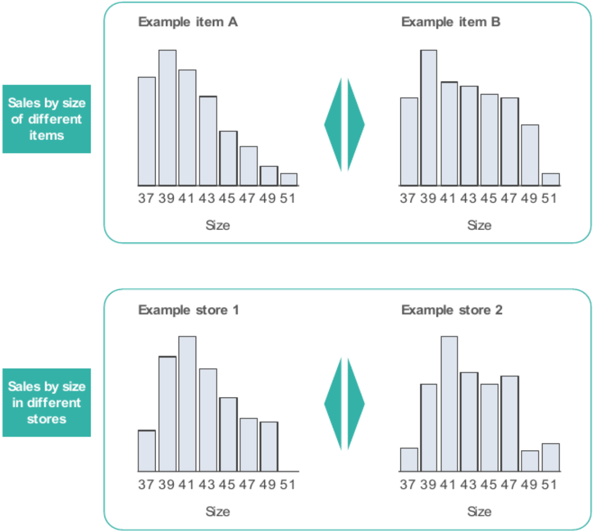
Including Manager Input on Customers and Local Trends:

1. Idea: providing each store manager with a **non-monetary budget** that they could use to accept or replace replenishment items that Miroglio's HQ proposed.
2. Garelli explained: "Our Head of Buying has a budget that she uses to decide how many blue jeans or black blouses to buy based on a demand forecast. However, she uses her judgment to modify these quantities from the demand forecast. We can offer a similar budget to store managers and effectively make them the buying heads of their own stores."
3. Head of Retail Carlo Tibaldi was very excited about this idea: approach would empower store managers and encourage them to embrace the analytical model.

Budget

Budget = store's potential sales in the next four weeks + safety stock of 20% - current stock

Example



Sales by size of different items
Example item A
37 39 41 43 45 47 49 51
Size
Example item B
37 39 41 43 45 47 49 51
Size
Sales by size in different stores
Example store 1
37 39 41 43 45 47 49 51
Size
Example store 2
37 39 41 43 45 47 49 51
Size

Example

	Item 1	Item 2	Item 3
Selling price (€)	100	200	150
Expected sales (units)	5	3	3
Stock (units)	2	4	2

Expected revenues: €1,550 = (5 x 100) + (3 x 200) + (3 x 150)
Potential revenue from stock: €1,300 = (2 x 100) + (4 x 200) + (2 x 150)
Extra coverage: €310 = 20% x 1,550
Budget: €560 = 1,550 + 310 – 1,300

Source: Evo Pricing.

Exhibit 3 Example of Recommended Actions

Store	Item	Stock (units)	Potential (units)	Difference (units)	Action
Store 1	Item 1	10	-1	9	Mandatory release
	Item 2	1	10	-9	Urgent replenishment
Store 2	Item 1	3	1	2	Recommended release
	Item 2	6	5	1	No action
Store 3	Item 1	7	8	-1	No action
	Item 2	2	5	-3	Replenishment

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\implies$ more frequent data $\implies$ better ML performance

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\implies$ more frequent data $\implies$ better ML performance
- Specify business problems by formulating them as decisions

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\Longrightarrow$ more frequent data $\Longrightarrow$ better ML performance
- Specify business problems by formulating them as decisions
    - Check if non-ML approaches can be used (benchmark)

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\implies$ more frequent data $\implies$ better ML performance
- Specify business problems by formulating them as decisions
  - Check if non-ML approaches can be used (benchmark)
  - Convert to ML (prediction) problems.

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\Longrightarrow$ more frequent data $\Longrightarrow$ better ML performance
- Specify business problems by formulating them as decisions
  - Check if non-ML approaches can be used (benchmark)
  - Convert to ML (prediction) problems.
  - Source of data: Unstructured (images) may be useful here.

Takeaways - 1

- Fashion is a challenging market to make predictions for several reasons. These factors are less relevant for some industries like groceries.
- Fast moving $\implies$ more frequent data $\implies$ better ML performance
- Specify business problems by formulating them as decisions
  - Check if non-ML approaches can be used (benchmark)
  - Convert to ML (prediction) problems.
  - Source of data: Unstructured (images) may be useful here.
- Challenge here is that important data may NOT be available to ML system

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations
  - Challenging to capture that data because human needs to be incentivized

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations
  - Challenging to capture that data because human needs to be incentivized
  - Verifying human input may not be possible.

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations
  - Challenging to capture that data because human needs to be incentivized
  - Verifying human input may not be possible.
  - Need to understand and design incentives carefully. Build in flexibility for humans to leverage their local knowledge.

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations
    - Challenging to capture that data because human needs to be incentivized
    - Verifying human input may not be possible.
    - Need to understand and design incentives carefully. Build in flexibility for humans to leverage their local knowledge.

Takeaways - 2

- Human may have access to that data, likely unstructred based on conversations
  - Challenging to capture that data because human needs to be incentivized
  - Verifying human input may not be possible.
  - Need to understand and design incentives carefully. Build in flexibility for humans to leverage their local knowledge.

Bottomline

AI challenges are embedded in a larger socio-economic context. Need to understand the ecosystem and design everything around the AI system.