Factorization Machines, Visual Analytics, and Personalized Marketing
Suneel Grover, Advisory Solutions Architect, SAS Institute Inc. and Professorial Lecturer, The George Washington University (Department of Decision Science - School of Business)

ABSTRACT
Competition in customer experience management has never been as challenging as it is now. Customers spend more money in aggregate, but less per brand. The average size of a single purchase has decreased, partly because competitive offers are just one click away. Predicting offer relevance to potential (and existing) customers plays a key role in segmentation strategies, increasing macro- and micro-conversion rates, and the average order size. This session (and the associated white paper) covers the following topics: factorization machines and how they support personalized marketing; how SAS® Visual Data Mining and Machine Learning, SAS® Decision Manager, and SAS® Customer Intelligence 360 support building and deploying factorization machines with digital experiences; and a step-by-step demonstration and business use case for the sas.com bookstore.

INTRODUCTION
Why does your organization’s website or mobile app exist? What are you hoping to accomplish with your business by being digital? What are the most important priorities for your digital presence?

Here are three goals that most organizations share:
- Sell more stuff.
- Make marketing more effective.
- Delight customers.

Goals are specific strategies leveraged to accomplish business objectives. Let’s walk through one of the examples cited above.

What does sell more stuff actually mean?
- Do x (more cross-sell or upsell efforts).
- Improve y (increase conversion rates).
- Reduce z (reduce abandon cart rates).

Every brand offers a digital experience for a reason (for example, if you’ve ever used Amazon or Netflix, you’ve experienced the value of recommendation systems firsthand). These sophisticated systems identify recommendations autonomously for individual users based on past purchases and searches, as well as other behaviors. Customers get algorithmically produced recommendations on additional offerings that are intended to be relevant, valued, and helpful. Consumers can use recommendations to do the following:
- Find things that are interesting or useful.
- Narrow a set of choices.
Marketers can benefit from recommendation systems to enhance offers that proactively build better customer relationships, retention, and sales. For example, organizations typically realize the following:

- Higher engagement, click-through, and conversion rates.
- New opportunities for promotion, persuasion, and profitability.
- Deeper knowledge about customers.

Recommendation systems aim to address the customer-centric paradigm that considers the different actions that a brand can take for a specific individual and decides on the “best” offering (such as one or multiple products). The goal is to produce an offer or proposition determined by the customer’s interests, and the organization’s objectives and policies. This is well suited for the following:

- Inbound real-time interactions, like websites, mobile apps, or call centers.
- Outbound scheduled or triggered interactions, like email or SMS text messages.

How are recommendation systems different from other forms of analytically-driven one-to-one marketing\(^1\) (excluding rule-driven targeting)? Let’s address this by defining the following three tiers differentiated by data availability:

1. A/B, multi-arm bandit, or multivariate testing\(^2\) – An extremely popular flavor of optimizing a digital experience that requires little data about the visitor to get started.

2. Recommendations – Interactions can be grouped and compared, allowing you to identify aspects of behavior that act as proxies for interests. Light to moderate amounts of data are required.

3. Propensity modeling\(^3\) – Predictive targeting based on probabilistic likelihoods using machine learning. Typically requires more data, such as behavior, history, location, demographics, and psychographics.

Before diving into the nuts and bolts of analytical approaches for recommendation system development, let’s take a moment to consider the value proposition of micro-moment marketing\(^4\).

The philosophy behind micro-moment marketing is that in the world today, consumers are bombarded by content, ads, offers, emails, texts, tweets, push notifications, and everything else imaginable. The industry has reached a point of “content shock”\(^5\) where consumers cannot digest much more content. Whether you are walking down a city street, sitting in a coffee shop, attending a sporting event, or vacationing at a resort, look around. What do you see? It’s stunning to observe the number of eyeballs viewing digital screens. In parallel, it isn’t a surprise that marketers venture to where audiences engage. The question is the following:
Do marketers need to adapt to capture the attention of consumers?

It's important to objectively realize that your brand or product isn't the center of your consumer's world. In fact, most marketing-centric content is perceived as an interruption to a customer-oriented experience. The key of micro-moment marketing is to embrace the notion that a few seconds exist to capture the attention of your target. Within that tiny window, brands are challenged to convey a communication that resonates.

Ultimately, the customer journey is just an amalgamation of micro-moments across channels, devices, and varying flavors of intent. Sometimes these moments will be focused on content consumption, raising awareness, or seeking service, while other interactions are more purchase-driven.

Pivoting back to recommendation systems, the main intent here is to provide the following:

- Align with key business imperatives.
- Deliver through customer-interaction channels.
- Inform targeted offers.
- Shape multichannel conversations.

Recommendation systems are evolving, and the topic of predictive analytics, machine learning, and artificial intelligence is the headline. At the center of the conversation is the desire to embed analytics with rule-based constraints into customer experiences to provide the optimal treatment, one step, or decision, at a time. Brands desire integrating analytical scoring to improve arbitration processes for offer propositions during real-time interactions, or when triggering communications.

**USER-CENTRIC VERSUS ITEM-CENTRIC FILTERING**

Marketers can use two flavors of algorithmic approaches to recommendations.

User-centric collaborative filtering is a very popular way of recommending products or content. It involves finding general similarities in behavior or characteristics between users then identifies specific differences between those linked users to make the case for sales opportunities.

For example, you might identify a group of customers who have purchased a common set of offerings. If 70 percent of a segment also have Product X, then you have a potential cross sell item for the 30 percent that have not bought it. The strength of the prediction for X (the filtering part) to a user depends on how similar that user is in behavior to others in the group (the collaborative part).
Item-centric (or content-based) filtering is used to recommend other products based on associations. This is helpful when you don’t know anything about the user other than they are interested in a product in their current web or mobile session. Product X is associated with Products Y and Z. A popular example of this approach is called association rule learning, which is a rule-based machine learning method for discovering relationships between variables. The association method is intended to identify strong rules discovered in data using measures of interestingness (support, confidence, and lift).

For example, the rule: \{Cavaliers, Warriors\} → \{Wizards\} found in the ticketing sales data of an NBA professional basketball organization generates a lift value of 3, indicating that if a fan has purchased tickets to see their local team play the Cavaliers and Warriors, they are three times as likely to have also bought a ticket to see the Wizards. Such information can be used as the basis for decisions about digital marketing activities such as promotional pricing, or targeted product recommendations.

MEASURING CONSUMER INTEREST IN PRODUCTS AND SERVICES

Recommendation systems require more than just building statistical models. Measurement and deployment requires the following:

1. Collect data about users, items, and profiles.
2. Prepare, aggregate, and filter data for specific types of algorithms.
3. Run an algorithm or algorithms and produce actionable scoring.
4. Deliver personalized recommendations.
Before you do any analysis, you need data. To enable an algorithmic recommendation model from a user's behavior, a distinction is often made between explicit and implicit forms of data collection.

Examples of explicit data collection include asking a user to do the following:

- Rate an item on a sliding scale.
- Rank a collection of items from favorite to least favorite.
- Add to cart, add to wish list, and posting to social media
- Purchases

Examples of implicit data collection include the following:

- Observing the items that a user views on a website or app.
- Search for terms or brands.
- Analyzing item/user viewing times.
- Session duration, frequency of sessions, and homogeneity of product views (within or across categories)

A decision needs to be made here. For example, your brand selects an implicit data collection method to capture a prospect’s interest in a specific product, offering, or piece of content. The simplest method is to measure whether a user views a product, and how many times. This can then be aggregated across website visits and/or app sessions. For each user and product, you end up with a count.

In reality, data is often highly skewed, with many visitors who viewed one product, or none. Plus, brands with multiple offerings might find that numerous products are never viewed. This is an example of the sparsity problem, and even the most active users of a website or app will have only viewed a small subset of the overall offerings.
Here are a few challenges every recommendation system must overcome:

**Scalability**: As a recommendations database grows, the performance decreases. It is beneficial to try to make systems that can handle large amounts of data and produce accurate recommendations quickly. There will always be a trade-off between performance and prediction accuracy; however, recent innovations in the parallelization of algorithm execution is addressing the challenge.

**Cold-start problem**: The problem appears at early stages of a recommender system’s life cycle, or when a new user or item is added to the system. If there is little information about users or items, the collaborative filtering will behave poorly.

**Sparsity**: It is common in digital marketing for people to view, click, or purchase relatively few items compared with the total number of items available. This leads to a sparse users-items representation matrix and, therefore, inability to identify neighbors or derive common behavior patterns. The result is low-quality recommendations. This problem is addressed in latent factor modeling algorithms that use dimensionality reduction on items and users, resulting in the identification of common behavior patterns within a reduced dimensional space.

**INTRODUCTION TO FACTORIZATION MACHINES**

According to Jorge Silva and Raymond E. Wright’s paper from SAS® Global Forum 2017, factorization models, which include factorization machines as a special case, are:

> "Broad class of models popular in statistics and machine learning. For example, principal component analysis is a well known type of factorization model that has long been a staple of dimensionality reduction. Matrix factorization has been widely used in text analysis and recommender systems. More recently, Rendle (2010, 2012) has proposed factorization machines for recommender systems and click-through rate prediction. Factorization machines are a powerful model that significantly extends matrix factorization."

In other words, a factorization machine is a predictive model that combines features of a support vector machine and matrix factorization. By modeling all variable interactions with factorized parameters, factorization machines can handle large data with high volumes of missing values (that is, sparse data), as well as trained in linear time.

Factorization models have attracted a lot of research in the fields of intelligent information systems and machine learning. They have shown excellent prediction capabilities in several important applications, such as personalized marketing and recommendation engines.

**HOW DO FACTORIZATION MACHINES SUPPORT PERSONALIZED MARKETING?**

The algorithm aims to learn consumer preferences to recommend items such as movies, books, songs, or other types of offerings. The purpose is to predict how a user would react to a set of items, and then recommend items that the user is likely to prefer the most. Whether in real time while customers browse, or through outbound marketing later, several things need to happen.
1. At first, while learning about your users' tastes and preferences, recommendations can be based on item attributes alone.

2. Over time (and with enough data), factorization machines enable useful analysis to deliver more meaningful recommendations.

3. Other users’ inputs can improve the results, qualifying why the system should be retrained periodically.

SAS Customer Intelligence 360 and SAS® Viya® provide an end-to-end solution, and the platform consists of the following:

- A scalable front end that records user interactions to collect, prepare, and store aggregated data that is analytically ready.
- Machine learning capabilities.
- Algorithms that can analyze and create relevant recommendations.
- Storage of user scoring that can be used by the deployment front end, in real time or later, based on the timeliness requirements for interactions.

**USE CASE: SAS BOOKSTORE**

In this example, you will see how the SAS bookstore can use factorization machines to compute relevant recommendations and present them to customers who are browsing the brand’s website. With countless titles available, predicting how customers would rate books enables a personalization strategy to prioritize showing specific book titles that they are likely to be interested in. For example, by knowing that customers are more likely to purchase a book with a rating higher than 4.5; therefore, we predict the rating and recommend books with ratings above 4.5.

SAS Customer Intelligence 360 can collect data about users based on the following:

- Implicit behavior.
- Explicit input.

For a bookstore, it is common for a subset of users to leave ratings. Outside of this specific use case, other proxies such as pageviews or screen views could serve as data inputs. However, in this example, users and items share a data trail that analysts can use to form a data-table matrix for modeling. Suppose you have user ratings of 1-5 where each value represents a rating given by the user.
We observe from the table that some of the ratings are missing, and we would like to predict these. Ultimately, the following is our end goal:

Notice that you can closely recreate the existing ratings, as well as get an approximation to the missing values. The intuition behind using matrix factorization to solve this problem is that there should be some latent features that determine how a user rates a book. Based on your digital property's traffic volumes, this matrix can become very large, because there can be millions of users and product items. Moreover, it is sparsely observed, because typically only a small fraction of historical ratings is available for analysis.
Users and items are characterized by their respective k-dimensional factor vectors. Revisiting Jorge Silva and Raymond E. Wright’s research effort referenced earlier:

“You can overcome these challenges by factorizing the matrix into lower-dimensional user and item factors, which can be used to predict new ratings. For recommender systems, the input vector is typically constructed using binary indicator variables for user u and item i.”

\[
x = (0, \ldots, 0, 1, 0, \ldots, 0, 0, \ldots, 0, 1, 0, \ldots, 0)
\]

\[|U|\quad |I|\]

Figure 5. Matrix Factorization for a Recommender System. Users and Items Are Characterized by their Respective K-dimensional Factor Vectors.

Figure 6. Input Vector for Recommender Systems

The factorization machine model is then equivalent to the following equation for predicting new ratings:
\[ \hat{y}(x) = \hat{y}(u,i) = w_0 + w_u + w_i + \sum_{f=1}^{k} v_{uf} v_{if} \]

**Figure 7. Factorization Machine Equation**

Moving on, because this example draws on data based on visitor interactions with the SAS bookstore, the marketing strategy is to proactively offer recommendations of books users didn't click on explicitly but might like. Our analysis data set in support of this objective has four columns captured and provided from SAS Customer Intelligence 360: Book ID, Visitor ID, Book Rating, and Time ID.

**Figure 8. Input Data Set Attributes**

Let's produce rating predictions for every specified Visitor ID and Book ID. Users select the factorization machine analysis object, drop it into the work space (see figure below), assign attributes to roles, and the model executes to provide results.
Using the FACTMAC\textsuperscript{12} procedure, you can call the action to implement the factorization machine model. It can be used to read and write data in distributed form to perform factorization in parallel by making full use of multi-core computers or distributed computing environments.
Here is an introductory video of building factorization machine visual models in SAS. For those readers who prefer programmatic approaches, view this.

The FACTMAC procedure estimates factors for each of the predictors (Book ID and Visitor ID), in addition to estimating a global and a level bias. After specifying the target variable (Book Rating), the procedure computes the biases and factors by using the stochastic gradient descent (SGD) algorithm, which minimizes the root mean square error (RMSE) criterion. In this method, each iteration attempts to reduce the RMSE. The SGD algorithm proceeds until the maximum number of iterations is reached, visually exemplified by the loss function within the iteration plot.

As shown in the image below, analysts have many optimization modeling options available. For example, the learn-step parameter controls how fast the stochastic gradient descent solver learns. Smaller values increase accuracy but might require a larger number of iterations to reach a good solution. Auto tuning is supported in selecting optimal values for factor count, maximum iterations, and learn-step. Lastly, to ensure model stability, data partitioning for table stratification is available for training, validation, and testing.
Further demystification of the model includes the scored response and assessment plots. The scored response visualization shows the distribution of the model’s predicted book ratings. The assessment graphic shows the overall performance of the model by comparing the predicted and observed book ratings. Assuming we are happy with this model, the next step is to make the rating scores available for deployment.

The output of this selection enables an analyst to produce the following data view.
Each user now has a predicted rating score for every product available in the SAS bookstore. And every visitor can have a rank-ordered set of scores associated with their profile. Rest assured, there are different approaches in making analytically scored data available within (and outside) of the SAS platform beyond a table. Here are some options that analysts have in helping their marketing teams:

- Scoring identifiable and anonymous traffic.
- Publish models in batch by calling the model through SAS code, Python code, or REST APIs.¹⁶
- Publish models in real time by calling the model through REST APIs using SAS® Micro Analytic Service.¹⁷

Specific to the usage of recommendation systems within marketing, SAS Micro Analytic Service is a powerful mechanism. For example, it can be called as a web application with a REST interface by SAS and other client applications. Envision a scenario where a visitor clicks on your website or mobile app, meets an event definition, and a factorization machine model runs to provide a fresh recommendation score to personalize the next page or screen of that digital experience. The REST interface (known as the SAS micro analytic score service) provides easy integration with client applications and adds persistence and clustering for scalability and high availability.
For more information about publishing and managing models in production, check out this article.18

INTELLIGENT DECISIONING WITHIN SAS CUSTOMER INTELLIGENCE 360

A typical day brings countless business decisions that affect everything from profitability to customer experience. What is a reasonable price point? Which audience segments should I personalize offers for? When should I recommend specific content earlier in a customer journey?

Daily decisions like these can alter the trajectory of a brand’s business. And while one bad move might not seem detrimental, hundreds or thousands of such operational decisions can be. So, it’s important that each decision is made with the best, most accurate information—while remaining consistent with organizational policy.

The output of the factorization machine model leads to scoring that identifies the best offer (or book) to deliver within a consumer touchpoint. The decision can be made in real time for inbound interactions and triggered for outbound communications.

Figure 14. SAS Customer Intelligence 360 - Decisioning Layer

Let’s walk through this. You have a prospect with a high recommendation score for a specific offer, that engages through social channels, but is under the age of 18. Your organization’s policy is not to advertise to individuals below a specific age threshold, regardless of how high the analytical score is. The decisioning layer of SAS Customer Intelligence 360 allows you to combine analytical models, business rule sets, and conditional logic into customer treatments and publish the decisions for orchestration.

Rule sets capture the logic of business decisions by allowing users to codify the decision-making process used by your organization. The rules make the decision-making process...
transparent and adaptable, enabling brands to respond quickly to new information about customers, segments, and markets.

Returning to our SAS bookstore use case, assume that I have been interacting with the brand’s owned digital properties across a visitor’s journey. After interacting with a call-to-action on a mobile app, I am redirected to a website and targeted with analytically recommended content.

The question is how did I configure event monitoring for that specific type of interaction with the mobile app? Within the user interface of SAS Customer Intelligence 360, there are a variety of interaction types that can be captured.

1. Select Event Type

Use events to identify how users interact with your content.

- **Add to Cart**
  Define an event when a user adds a product to cart

- **Cart View**
  Define an event when a user views a cart

- **Check-out View**
  Define an event when a user views a check-out

- **Click**
  Define an event when a page element is clicked

- **External**
  Define an event that interacts with an external application

- **Form Submit**
  Define an event when a user submits a form

- **Mobile**
  Define an event when a user interacts with a mobile application

- **Page View**
  Define an event when a user views a page

- **Product View**
  Define an event when a user views a product

- **Purchase View**
  Define an event when a user views a purchase

Figure 15. SAS Customer Intelligence 360 - Event Capture
For this example, we selected the mobile event type. After configuration is complete, it is that interaction on the mobile app that will generate an event in SAS Customer Intelligence 360 and trigger the targeting on the website.

Now that we can observe every instance when that interaction occurs, we want SAS Customer Intelligence 360 to contact SAS Decision Manager on SAS Viya in real time for every visitor journey that meets the event definition. To do so, an agent is established to make the API connection.

![Figure 16. SAS Customer Intelligence 360 - Authoring Agent Credentials](image)

Once the credentials are set, the association of the visitor interactions with the mobile app event and SAS Decision Manager can be defined.
The value of unifying these two technologies within the SAS platform creates the following powerful benefits:

- SAS Customer Intelligence 360 can be used to stream real-time behavioral events to SAS Decision Manager.
- Real-time events from digital touchpoints can trigger decisions in SAS Decision Manager. These decisions can be used with other data sources such as CRM, demographics, transaction history, third-party data, and so on.
- Decisions executed by SAS Decision Manager can manage and execute analytic models, business rules, and conditional logic to return a recommended course of action.

Let’s break this down, one step at a time. When the mobile app event occurs, SAS Decision Manager will use the SAS Micro Analytic Service to perform the following:

1. Query the latest CRM (customer relationship management) values for this specific customer regarding their education, income, and age from a database that SAS Viya is connected to on-premises or in the cloud.
2. Immediately apply a business rule that offer personalization cannot be targeted at individuals aged under 18.
3. Run an auto-tuned, factorization machine model to produce a fresh recommendation score using digital experiential data captured and analytically prepared by SAS Customer Intelligence 360.
4. Apply conditional logic using different combinations of additional CRM business rules.

![Figure 20. SAS Customer Intelligence 360 - Conditional Logic for Defining Segments](image)

5. Based on the customer's assignment, produce four actionable segments with additional context beyond just an analytical score, and inform SAS Customer Intelligence 360 how to personalize the offer on the next page of the current web visit.

The SAS Micro Analytic Service is part of SAS Decision Manager. It is a powerful feature designed to execute analytical models and business rules against the latest data from online channels, combined with data from operational databases and other data sources. This entire process executes in milliseconds and works with SAS Customer Intelligence 360 to increase personalization precision without disrupting the customer’s digital experience.

Whether the intent is to display a single recommendation like the following:
Figure 21. SAS Customer Intelligence 360 - Single Product Item Targeting

Or if the intent is to allocate multiple recommendations at once like the following:

Figure 22. SAS Customer Intelligence 360 - Multi-product Item Targeting
The beauty is all the execution ideas are options for the marketer to consider. When it comes to targeting audiences for web personalization, mobile cross-sell opportunities, or email acquisition programs, this decisioning output is crucial in including (or excluding) individuals for recommendations.

**Figure 23. SAS Customer Intelligence 360 - Multi-channel Delivery Capabilities**

The marketing strategy will lead with a recommendation concerning the best offer (or offers) to deliver within the touchpoint. The decision will select only the highest ranked proposition or propositions out of the possible book title offers, depending on the individual's predicted score or scores, their eligibility, and other relevant criteria. Here is an example of how a marketer would set up the nested targeting logic within SAS Customer Intelligence 360:

**Figure 24. SAS Customer Intelligence 360 - Audience Targeting**
CONCLUSION

SAS recognizes there are different user segments that leverage our platform, and this white paper focused on how analysts who prefer to build custom analytical recommendation models in support of the marketer’s agenda for delivering customer personalization. Although not covered, for readers interested in automated recommendation systems embedded within SAS Customer Intelligence 360, read this article\textsuperscript{19}.

![Figure 25. SAS Customer Intelligence 360 - User Personas](image)

At the heart of this white paper is the FACTMAC procedure, which enables you to solve a variety of tasks, from recommendations to predictive modeling, all of which involve sparse data. Thanks to a highly parallel optimization solver, PROC FACTMAC can handle very large data sets. This powerful and flexible method provides not only predictions but also meaningful factor representations that can give you insights into many types of business problems.

Although we took a tour of how SAS Customer Intelligence 360, SAS Visual Data Mining and Machine Learning, and SAS Decision Manager can work together, ultimately SAS wants to help marketers be effective through analytic techniques. Consumer preferences are difficult to predict. By exploiting the deep library of algorithms provided by SAS, intelligent decisioning features, and orchestration delivery, recommendations can automatically shape shift to meet the demands of the consumer and create brand relevancy through data-driven personalization.
REFERENCES


RECOMMENDED READING

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Suneel Grover, M.B.A., M.S.
Advisory Solutions Architect / Professorial Lecturer
SAS Institute, Inc. / The George Washington University, Washington D.C.
Email: suneel.grover@sas.com
LinkedIn: https://www.linkedin.com/in/suneelgrover
Twitter: @suneelgrover
Web: https://blogs.sas.com/content/author/suneelgrover/

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.