Executive Summary

Team 51

John Kiley, Brian Merrill,
Hemant Patel, Julia Rodd

cognoclick

Table of Contents

About Us

We at CognoClick know what it takes to build a successful recommendation engine. We have worked with clients across a wide array of industries and have helped each one exceed their sales goals through effective product recommendations. Using artificial intelligence and state-of-the-art natural language processing (NLP) techniques, we have helped clients make better use of their purchase history and product review data. Our proprietary solution allows clients to retain existing customers and attract new ones by offering highly relevant products.

"Our cognitive and predictive analytics will increase customer sales by identifying high relevance products to both new and existing customers."

About Our Target Company

Amazon.com Inc. (Amazon) is a worldwide e-commerce leader and technology company that offers various products and services, ranging from online shopping to video streaming and web services. With a mission "to be the Earth's most customer-centric company," Amazon has experienced great success (Amazon's global career site). In 2018, Amazon reported over 141 billion dollars in sales and over 232 billion dollars in earnings across U.S. and international markets (AMZN.O - Amazon.com, Inc. Profile). Looking beyond the numbers, Amazon has had a lasting impact on the online shopping experience. By carrying millions of products across a multitude of categories, Amazon is able to appeal to a wide variety of customers and distinguish itself from the competition through customer choice.

Problem Statement

Since its inception nearly 25 years ago, Amazon has been focused on delivering a world class customer experience. In a world where consumer choice is paramount, Amazon continues to need better ways to target new and existing customers by presenting them with products they are most likely to purchase. The key to solving this problem lies in one of Amazon's greatest data assets: **customer reviews**.

"Amazon product reviews are the most popular and trusted, and [customers] will go to Amazon for reviews even if they intend to purchase the product elsewhere" (Vega, 2017).

Amazon has a subset of customers who write reviews about their purchased products (reviewers). Reviewers tend to be more engaged with Amazon's platform, giving Amazon the ability to increase product exposure with this segment.

By mining historical reviews for information on product preference, Amazon can develop a more personalized product recommendation experience for each reviewer.

Solution

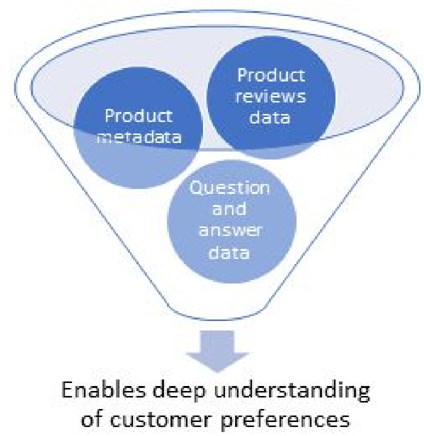
CognoClick will expedite Amazon's opportunity to drive business value by creating a more personalized product recommendation experience for its reviewers.

By delivering personalized and high relevance product recommendations to reviewers, CognoClick will help Amazon experience the many benefits of a recommendation engine, including increased product impressions and better insights into the needs of a key customer segment.

In making this investment and partnering with CognoClick, Amazon will experience an immediate revenue lift in the short term and significant return on investment over the long run.

CognoClick is leveraging 15 years of product review data from Amazon's electronics product category to build a revamped recommendation engine.

Through this process, the team is creating relevant features at a reviewer-level and a product-level. Additionally, the team is using NLP techniques to incorporate product review and text features. By starting small, CognoClick will gain the incremental feedback needed to prepare for scaling across Amazon's other product categories.



CognoClick's goals for the 10-week electronics recommendation engine POC are as follows:

- Assess different modeling approaches and build a recommendation engine that best recommends products to each reviewer
- Build a recommendation engine that incorporates both review text and metadata features
- Leverage NLP techniques to process review data and incorporate keywords as well as product and user features into the model
- Build and implement a model framework that is generalizable to Amazon's other product categories
- Create an interactive dashboard and mobile application that allows Amazon to monitor results and business impact
- Define additional opportunities for the current recommendation engine that will be addressed in a Phase 2 implementation

Project Status

The CognoClick team is on track to complete the 10-week POC. The team continues to remain in close communication and has been committed to delivering above and beyond project goals.

The only remaining deliverable is the CEO presentation, and there are no risk items to report. Figure 1 recaps the project timeline and status.

Figure 1: Project Timeline and Status

Recommendation Engine Background

The CognoClick team performed extensive research to determine how best to build Amazon's revamped electronics recommendation engine that provide the best success for high relevance.

"The basic task of a recommender system is to suggest relevant items to users, based on their opinions, context, and behavior" (McAuley et. al, 2015).

Overall, there are three industry-standard methods. Figure 2 provides an illustrative overview of these methods.

1. **Content-based filtering method**, which recommends items to a user based on historical reviews or interactions by that user.

2. **Collaborative filtering method**, which recommends items to a user based on historical reviews or interactions by other similar users/items.

3. **Hybrid method**, which recommends items to a user by combining results from both content-based and collaborative filtering methods.

Cold start is a common problem for recommendation systems and was considered a key risk that weighed heavily on CognoClick's ultimate modeling approach decision.

The cold start problem is an issue with collaborative filtering models. The model is unable to recommend a new product or generate recommendations for new users, since this kind of model relies on past interactions.

Per Falk (2019), the best and most robust recommendation engine method against the cold start problem is content-based filtering. Therefore, the CognoClick team knew it had to incorporate a content-based component in order to set Amazon up for success in creating recommendations for new users.

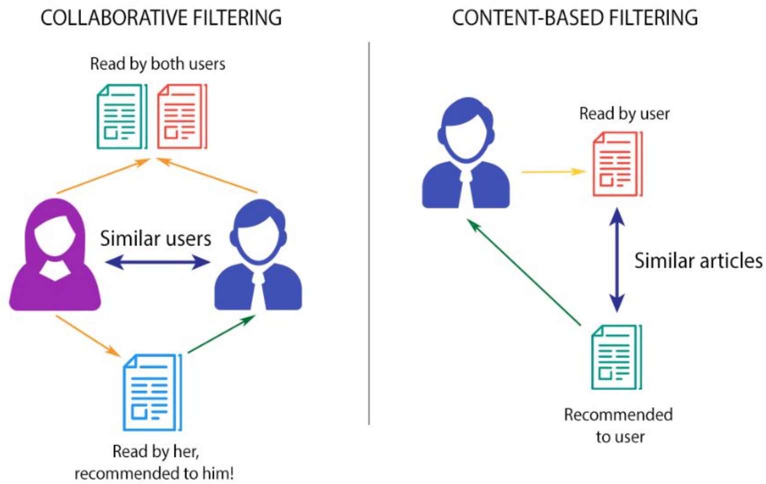


Figure 2: Collaborative and content-based filtering (Liao, 2018)

Modeling Approach

Given the three approaches outlined above of content-based filtering, collaborative filtering, and a hybrid method, **the CognoClick team chose to implement a hybrid method**, as it expects this method will perform better than either a content-based or collaborative filtering method alone.

Moreover, a hybrid method will not only mitigate cold start problems but will offer Amazon the scalability it needs for its ever-changing reviewer and product base.

Lastly, with a hybrid model, there is additional flexibility to incorporate diverse features, including features around users, products, and review text, a core goal of this POC.

Understanding the opportunity to test and leverage multiple models, the CognoClick team identified three independent models as feasible for the POC. The three models that were selected for CognoClick's recommendation engine include:

1. **Baseline model**, using product ratings data only.
2. **Review text model**, using features derived from review text.
3. **Deep learning model**, using metadata and text features from both reviewers and products.

The modeling was performed in two phases which are called the **Phase 1 Build** and **Phase 2 Build** for the sake of simplicity. The Phase 1 Build involved creating prototypes for all three models and then scaling each model if and when possible.

During the Phase 2 Build, the CognoClick team leveraged the learnings, outputs, and findings from the Phase 1 Build to generate model enhancements. Figure 3 illustrates these two phases and provides a preview into how the models were leveraged together.

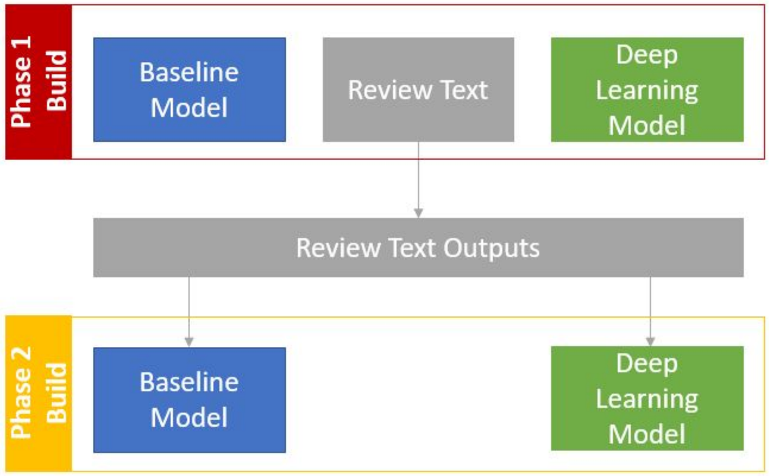


Figure 3: CognoClick Model-Building Strategy

There is one specific challenge that should be mentioned: the CognoClick team experienced memory issues and/or long model training time due to the large data set size (in terms of number of records and complexity of features) chosen for the POC.

To help speed model development across all models and generate results in the Phase 1 Build, the team used a subset of the electronics data from the Camera & Photo category. In addition, the team took full advantage of cloud resources, specifically through Google Cloud Platform, to enable faster speed of delivery.

Discovery and Results

Data Overview

In order to successfully build the recommendation engine, the CognoClick team carefully prepared the data and created relevant metadata and text features for both reviewers and products.

The electronics product data is comprised of 1.6 million reviews from 192 thousand reviewers across 63 thousand products.

The CognoClick team spent time inspecting and understanding the data. During this process, the team was able to see that most of the reviews have a positive rating (4 or 5 rating on a 5-point scale). Figure 4 below supports this point by showing the average ratings for each reviewer and each product.

It is important to note that the average ratings for reviewers are more volatile than the average ratings for products. This result supports industry-standard observations that product tastes are thought to be more stable over time.

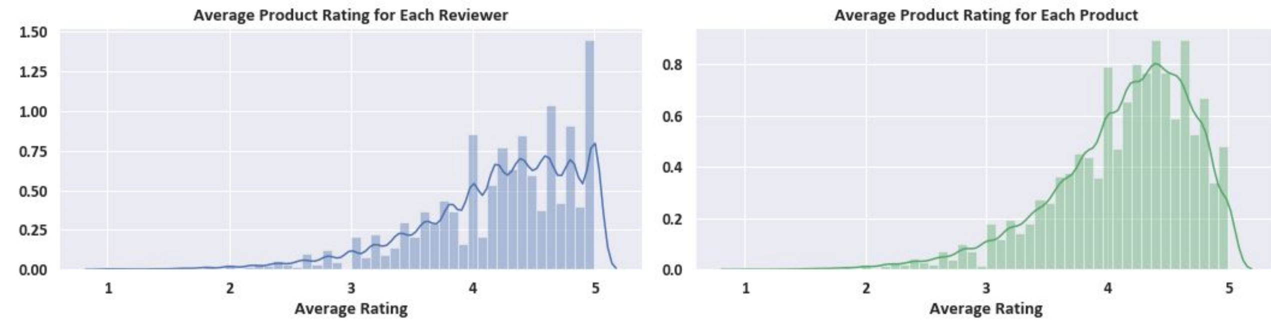


Figure 4: Average Ratings for Each Reviewer (left) and Products (right)

Results

Phase 1 Build

The **baseline model** was built leveraging the Scikit Learn matrix factorization for recommendation systems called Surprise. The model is intended to predict ratings for a user for products that have not yet been reviewed. If a prospective reviewer has never had a review the model would suffer from the inability to generate recommended reviews.

Therefore, the baseline model effectively leveraged sparse matrices to impute the missing values for each reviewer for each product via a process called matrix factorization. The prototype generated predicted ratings for a small number of users on all their unseen products in addition to creating recommendations on products that users had reviewed.

The **review text model** used the rich text data to find reviewer-reviewer and product-product similarity identifying cohorts of similar users and items. The review text model required significant natural language preprocessing and cleansing prior to passing the resulting output to k-nearest neighbors algorithm for clustering.

For computational effectiveness, ten clusters were selected. Both products and reviewers were clustered independently of each other resulting in a cluster assignment for every reviewer and a cluster assignment for every product. The cluster assignments were able to be leveraged by the baseline model during the Phase 2 Build.

One of the key deliverables in the Phase 1 Build was the term-frequency, inverse document frequency (TFIDF) matrices from the review text clusters. TFIDF generates a representation of both the frequency and the uniqueness of a word with a review and was able to serve as input into the deep learning model (Lane et al., 2019).

The **deep learning model** initially leveraged only the most basic data to generate product recommendations. The basic Phase 1 deep learning model lacked meaningful and detailed features but was able to run successfully on large volumes of data.

The model also showed significant promise on outperforming the baseline model in nearly every success metric, including computational performance and adaptability. Through these insights, the CognoClick team pivoted and divided further attention to enhancing the deep learning model.

Phase 2 Build

The focus for the **baseline model** in the Phase 2 Build was to generate predicted ratings for all users on all their unseen products. However, the baseline model failed to be executable on a desktop machine. The model was then lifted to a server on Google Cloud Platform (GCP) for execution. The machine was scaled beyond what many enterprises have available for data science problems, but still could not handle generating predictions.

At this point, the baseline model was revisited in the spirit of agile development. The model was successfully run on two of the review text user clusters, before exhausting all available GCP credits. All predicted ratings had a value of 5 (out of 5 stars), and the CognoClick team determined that this model was not valuable in delivering highly relevant recommendations. Consequently, additional work on the baseline model was ceased.

The **deep learning model** was also significantly expanded. The deep learning model was updated to leverage both metadata features and the term-frequency, inverse document frequency matrices created from the review text model that was produced as part of the Phase 1 Build.

This new input provided the deep learning model additional input to increase the number of features that were exposed to the model. Despite being the most complex in nature, the deep learning model ultimately had the best performance from a technology perspective, being able to be run on a local desktop computer.

Figure 5 captures a summary of the benefits and drawbacks of each of the models CongoClick considered en route to building its recommendation engine for Amazon.

To execute on its desired hybrid approach, CognoClick selected a best content-based and collaborative filtering model. Both of these models were created through a deep learning approach, as this method proved most effective at generating accurate and relevant recommendations.

Approach	Strengths	Gaps	Versatility to Integrate	Technology Demands
Baseline	Industry standard	Cold start	Limited	High
Review Text	Rich data source	Cold start	High	Medium
Deep Learning	Performance; Could build both content based and collaborative models	-	High	Relatively Low

Figure 5: Model Summary

Upon completion of the individual models, an **ensemble approach** was taken in an attempt to draw on the unique strengths of each model while providing coverage for the gaps of each model. The ensemble approach was limited to Cluster 0 and Cluster 1 due to the limited return results from the baseline model.

The ensemble model was a straightforward rules based ensemble. Figure 6 provides a high level overview of the model ensemble architecture. The product ratings were averaged across all models and rank ordered. The products with the highest average ratings were evaluated as the best recommendations provided by the ensemble model. This approach is logically consistent with how classification models operate with voting methods.

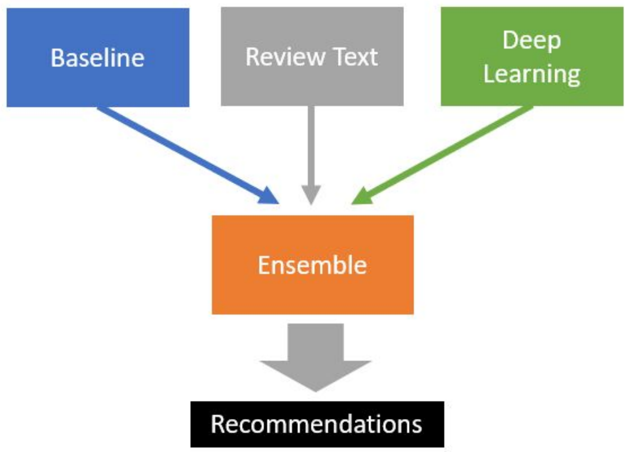


Figure 6: Model Ensemble Overview

Dashboard and Mobile Application

CognoClick created a dashboard and mobile application that allows Amazon to understand the product recommendations at a deeper level. Figure 7 provides a screenshot of the dashboard and mobile application. The dashboard and mobile application consists of two main components:

1. Overall product demand
2. A rank-ordered table containing recommendations for a given product or user

The overall product demand visual captures the top ten most popular products within the electronics product category. By hovering over each product, Amazon can glean additional context around a given product, such as the number of reviews, price, and its sub categories.

More importantly, CognoClick's dashboard and mobile application contains two tables that capture rank ordered recommendations at the product-level and user-level. The product recommendations illustrate the ten most similar products for a given product, while the user recommendations capture the ten most similar product recommendations that are most relevant for a user. Both tables capture additional metadata such as number of reviews, product price, average product rating, and product URL.

Dash was recommended by CognoClick for its extensibility and versatility. The code-intensive Dash application required significant upfront investment to build baseline capabilities. However, the upfront investment affords the application portability across web frameworks. The code base that was generated to support this initiative can be readily integrated into the existing Amazon enterprise web framework or ported to a new platform or system on the fly.

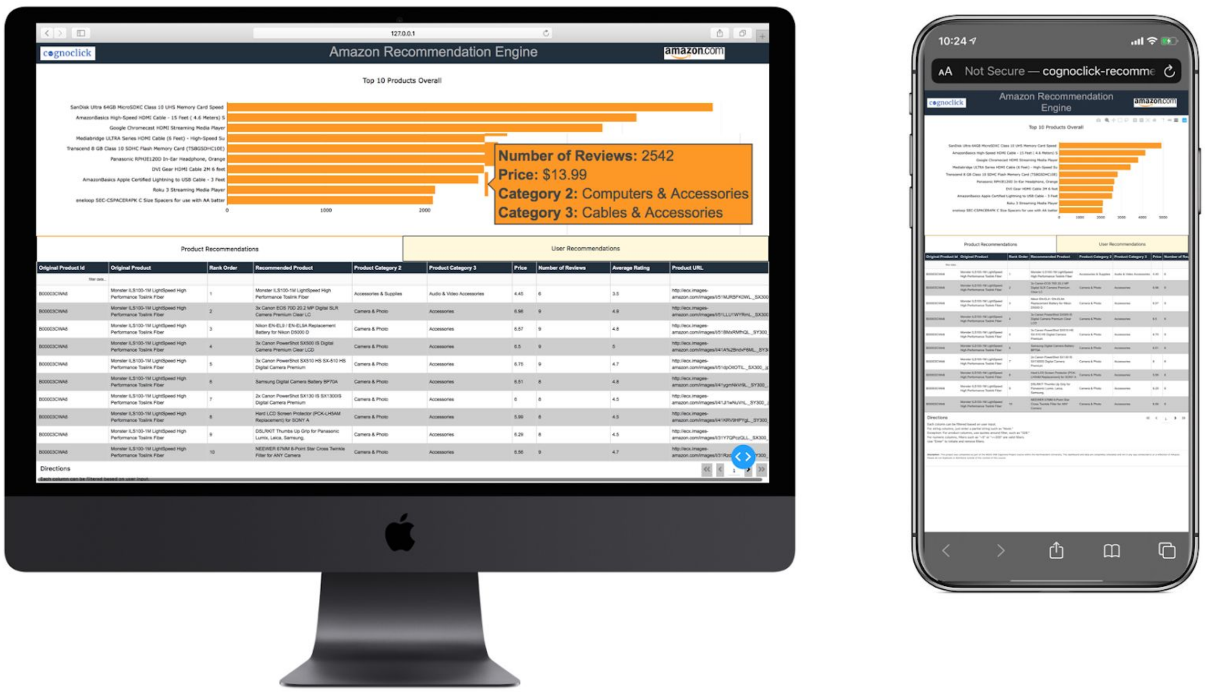


Figure 7: Dashboard (Left) | Mobile Application (Right)

Conclusions

A summary CognoClick's conclusions are outlined below. Full details can be found in CognoClick's final paper.

- An agile and incremental model development process allowed the team to continuously advance multiple models and generate insights and learnings along the way.
- A deep learning approach shows vast improvement over a baseline model and offers the ability to incorporate both metadata and text features in an extensible way.
- A hybrid model can utilize the diverse and feature-rich data to a much fuller extent than either content-based and collaborative filtering methods can alone.
- Review text features are extremely rich should be leveraged whenever possible. Using review text to create reviewer and product clusters provided the ability to identify similar reviewers and similar products as well as what is important to those reviewers.
- There is no shortage of opportunity to invest in model development. Amazon will need to leverage a suite of models as this approach offers the best opportunity to generate highly relevant product recommendations.

- Determining the relevancy of product recommendations is difficult without end user feedback. Interleaving is the preferred option to consider when evaluating and selecting final model(s), as end users are able to directly interact with and vote for the winning model(s).

- The electronics data is rich and therefore, computationally expensive. CongoClick leveraged the cloud to take advantage of its computing power and address memory issues. Amazon will need to do the same to support its revamped product recommendations.

Recommendations

The CognoClick team has put together a list of recommendations that Amazon should consider to take full advantage of the revamped electronics product recommendations. These recommendations are broken out into two phases:

1. **Phase 1 ("Building foundation")**: In this phase, Amazon implements the CognoClick electronics product recommendations and prepares to incorporate enhancements.

2. **Phase 2 ("Enhancing foundation")**: In this phase, Amazon incorporates enhancements into the CognoClick electronics product recommendations, monitors impact and results, and scales recommendations to other product categories.

Phase 1: Building Foundation

The objective of Phase 1 is to ensure that the completed work is implemented into production. Phase 1 establishes the critical foundational elements for all Phase 2 recommendations. Figure 8 below outlines the Phase 1 components.

The first step is to ensure that the infrastructure can support the computational requirements of CognoClick models. Amazon has world class technology with AWS in its portfolio of services, so this step is not expected to be an issue.

Amazon should then work to formally migrate the models to production. Once in production, the dashboard and mobile application can be utilized for managing results. Additional support may be required, gathering operational and behavioral elements from data that was not provided to CognoClick for the purposes of this POC.

Moving the model to production and tracking its success will ensure that Amazon's marketing, operations, and analytics teams are able to demonstrate the value of its investment in the revised recommendation engine.

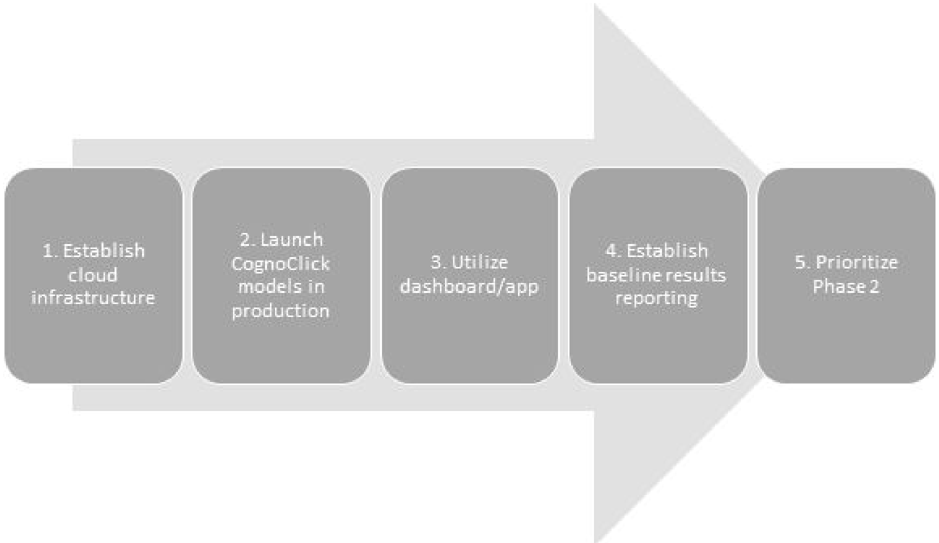


Figure 8: Phase 1 Recommendations

Phase 2: Enhancing Foundation

Good work begets more work. CognoClick's modeling framework is highly extendable given the right access to the Amazon operations data. To take full advantage of CognoClick's revamped recommendations, Amazon must implement a series of enhancements to ensure that the recommendations are taking advantage of all of the rich data elements available to Amazon, beyond what was provided as part of the POC, which should lead to continuous customer engagement and favorable bottom-line impacts. Examples of these data enhancements include web analytics data and other product data that is proprietary to Amazon. Figure 9 outlines these additional enhancements.

Figure 9: Phase 2 Recommendations

Once Amazon has enhanced its foundation for the electronics product category, it can then expand CognoClick's recommendations to other product categories using a similar approach. The model is generalizable and set up such that any other product categories that are passed to the model will generate recommendations.

CognoClick recommends that Amazon execute the models on a subcategory by subcategory basis to ensure that the increased scale from the additional data does not introduce additional computational complexity to the process.

The current model is set up such that it is designed to benefit one of Amazon's most valued resources: its reviewers. The learnings from these models can be extended to all users of Amazon's platform and, with the enhanced web analytics data, provide value to every visitor of the website.

The recommendations should be tested via A/B testing, or industry best practice interleaving. Testing on real customers will provide Amazon confidence and learnings that can be leveraged to continuously improve the models. The Dash dashboard and mobile application is extensible and can be adapted to the increased scope of the models.

Last, but certainly not least, a procedure for keeping the models up-to-date and highly-tuned should be established to ensure that they remain relevant and valuable assets over time. Retraining the models will be a requirement to protect Amazon's investment in the revamped recommendation engine.

References

Amazon's global career site. (n.d.). Retrieved October 4, 2019, from https://www.amazon.jobs/en/working/working-amazon.

AMZN.O - Amazon.com, Inc. Profile. (n.d.). Retrieved from https://www.reuters.com/companies/AMZN.O.

Dash User Guide and Documentation - Dash by Plotly. (n.d.). Retrieved from https://dash.plot.ly/.

Falk, K. (2019). *Practical recommender systems*. Shelter Island, NY: Manning Publications Company.

Lane, H., Howard, C., & Hapke, H. M. (2019). *Natural language processing in action: understanding, analyzing, and generating text with Python*. USA: Manning Publications.

Liao, K. (2018, November 19). Prototyping a Recommender System Step by Step Part 1: KNN Item-Based Collaborative Filtering. Retrieved from https://towardsdatascience.com/prototyping-a-recommender-system-step-by-step-part-1-knn-item-based-collaborative-filtering-637969614ea.

Mcauley, J., Pandey, R., & Leskovec, J. (2015). Inferring Networks of Substitutable and Complementary Products. Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD 15. doi: 10.1145/2783258.2783381

Vega, N. (2017, March 20). Here's why user reviews on sites like Amazon are such a big deal. Retrieved October 4, 2019, from https://www.businessinsider.com/amazon-reviews-greatly-impact-online-shopping-sales-2017-3.

Appendix

Code

The team used Github to store and manage all code for the project.

Here is a link to the repository: **https://github.com/dashpound/capstone**.

Dashboard and Mobile Application

The team created a separate repository to deploy the dashboard and mobile application via Heroku.

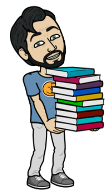
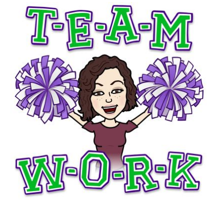
Here is a link to the final dashboard and mobile application: **https://cognoclick-recommendations.herokuapp.com/**.

Here is a link to the separate Github repository: **https://github.com/dashpound/review_dashboard**.

Project Team

CognoClick was born out of four Northwestern University MSDS students who have a passion for shopping, machine learning, and natural language processing. The team has strong technical and communication skills and is dedicated to delivering above and beyond project goals. Each team member also brings unique interdisciplinary skills that will help keep the project moving forward at all times. The project roles are summarized in the table below.

P = Primary S = Secondary * = All participating	**John Kiley**	**Brian Merrill**	**Hemant Patel**	**Julia Rodd**

				
Project Manager	S			P
Technical Writer	S	S	S	P
Programmer	P	S	P	S
Dashboard/Mobile App		P	S	
Oral Presentation	*	*	*	*

John Kiley is a Director of Data & Analytics at a financial services company and has a record of tackling challenging business problems and working in turn-around organizations. John's primary focus is on enabling better business outcomes through data driven decision making. John has several years experience leading machine learning, system conversion, data warehousing, and business intelligence projects. John will serve the role of programmer and will support project management and technical writing activities of the project.

Brian Merrill is a Managing Director at a Big 4 consulting firm providing data analytics services to clients to help solve their complex legal and compliance issues. With over 15 years experience developing applications, analyzing systems, data mining and performing data analysis, Brian brings significant insight and extensive client experience to CognoClick to help solve the difficult challenges of product filtering and recommendation. Brian will serve as the primary visualization expert as well as supporting the technical writing and development of the recommendation engine.

Hemant Patel is a Senior Manager at a performance marketing agency responsible for designing, developing, and implementing predictive modeling solutions for clients looking to optimize their marketing strategies. Hemant has an extensive background that includes experience in data mining, predictive modeling, and business analytics. Hemant will serve as a programmer as well as support technical writing and visualization efforts.

Julia Rodd is a Senior Data Scientist at a financial services company and has several years' experience leading and participating in analytics projects. She brings an ability to think strategically, technical knowhow, and a focus on driving for results. Julia will be serving the role of project manager and primary technical writer. She will also support the development work as needed.