aws SUMMIT

Build a Recommendation Engine on AWS Today

Yotam Yarden

Data Scientist, Amazon Web Services

Agenda

- **Recommendation Engine – Why?**
- Recommendation Engine – Common Techniques
- Introducing Amazon SageMaker
- Develop, Train & Deploy a Recommendation Engine in 15 minutes
- Customer use cases

Welcome to Amazon.com Books!

amazon.com
Earth's biggest bookstore

One million titles, consistently low prices.

(If you explore just one thing, make it our personal notification service. We think it's very cool!)

Spotlight! -- August 16th

These are the books we love, offered at Amazon.com low prices. The spotlight moves **EVERY** day so please come often.

One Million Titles

Search Amazon.com's million title catalog by author, subject, title, keyword, and more... Or take a look at the books we recommend in over 20 categories... Check out our customer reviews and the award winners from the Hugo and Nebula to the Pulitzer and Nobel... and bestsellers are 30% off the publishers list...

Eyes & Editors, a personal notification service

Like to know when that book you want comes out in paperback or when your favorite author releases a new title? Eyes, our tireless, automated search agent, will send you mail. Meanwhile, our human editors are busy previewing galleys and reading advance reviews. They can let you know when especially wonderful works are published in particular genres or subject areas. Come in, meet Eyes, and have it all explained.

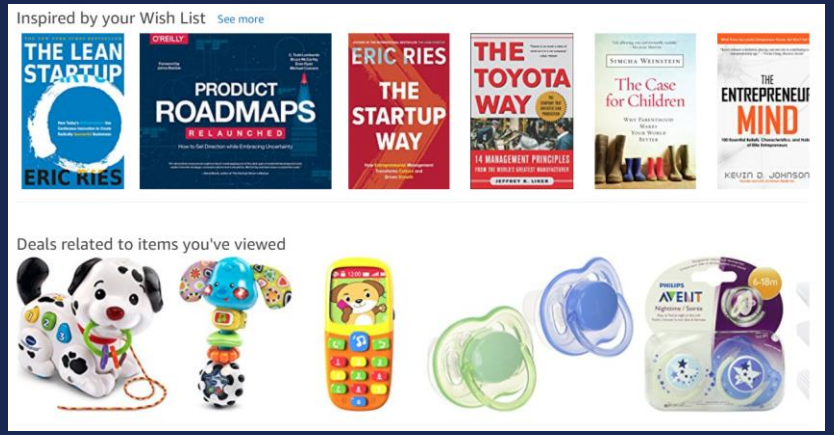
Your Account

Check the status of your orders or change the email address and password you have on file with us. Please note that you **do not** need an account to use the store. The first time you place an order, you will be given the opportunity to create an account.

Artificial Intelligence At Amazon (1995)

And today…

My Profile – amazon.de



My Profile – amazon.com

Motivation

- Personalize and enhance customer experience

- Different goals:
  - Increased time spent on a platform
  - Suggest complementary items
  - Customer satisfaction

Use Cases

Ecommerce:

- Amazon.com

Content:

- Movies (Netflix)
- Music (Amazon Music)
- Articles (The Global And Mail)

Finance:

- Services Recommendation
- Stocks buying / selling
- Relevant news and stock related data

Education:

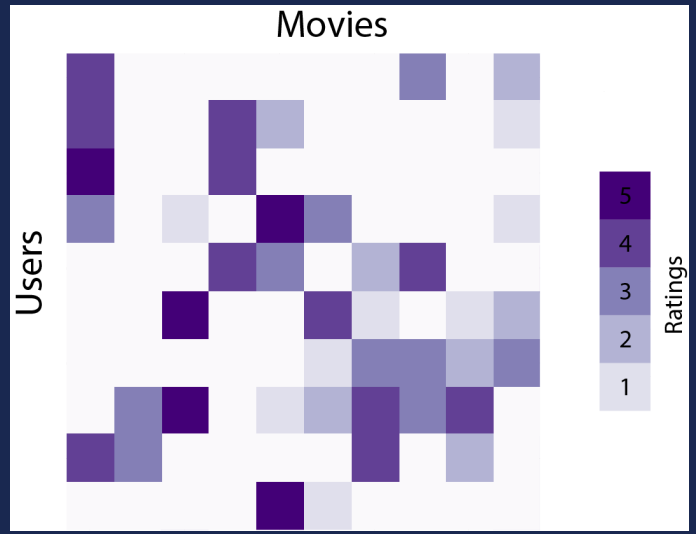
- Courses recommendations

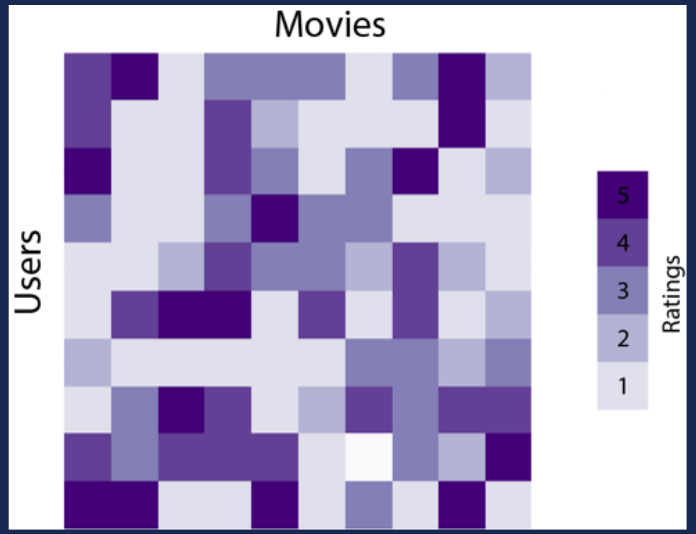
Legal:

- Similar cases

Agenda

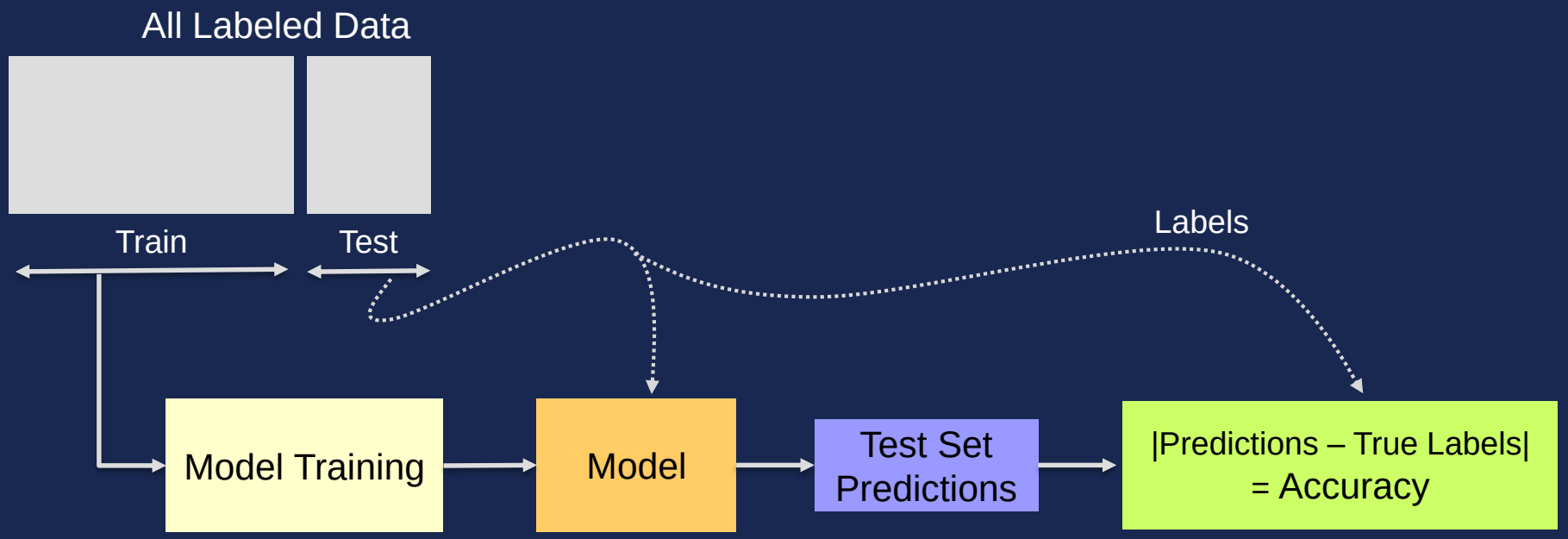
- Recommendation Engine – Why?
- **Recommendation Engine – Common Techniques**
- Introducing Amazon SageMaker
- Develop, Train & Deploy a Recommendation Engine in 15 minutes
- Customer Use Cases



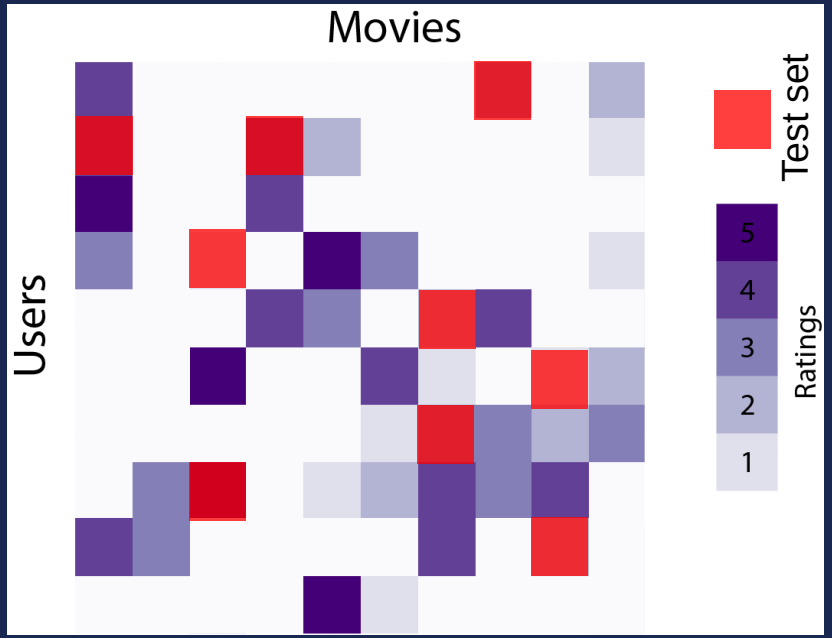


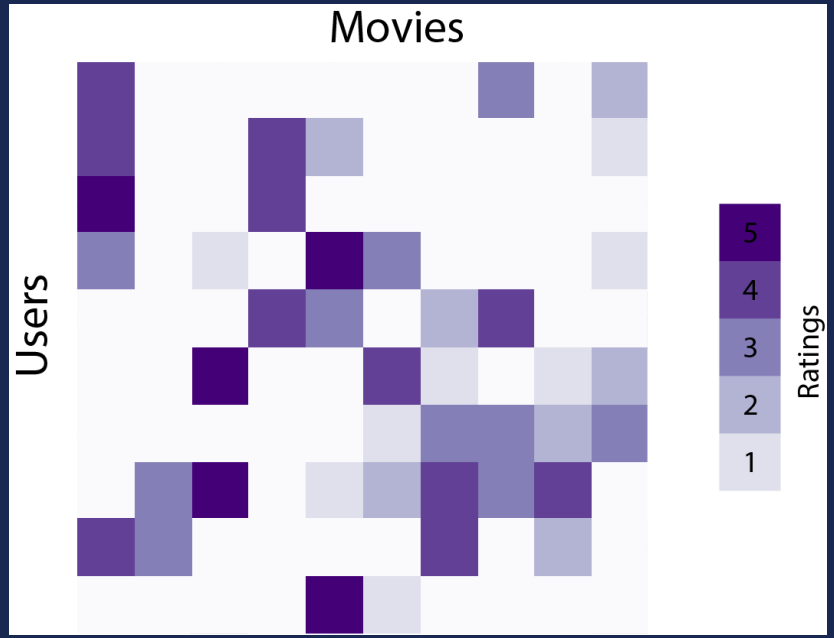
https://www.oreilly.com/ideas/deep-matrix-factorization-using-apache-mxnet?cmp=tw-data-na-article-engagement_sponsored+kibird

Supervised Machine Learning



Test / Validation





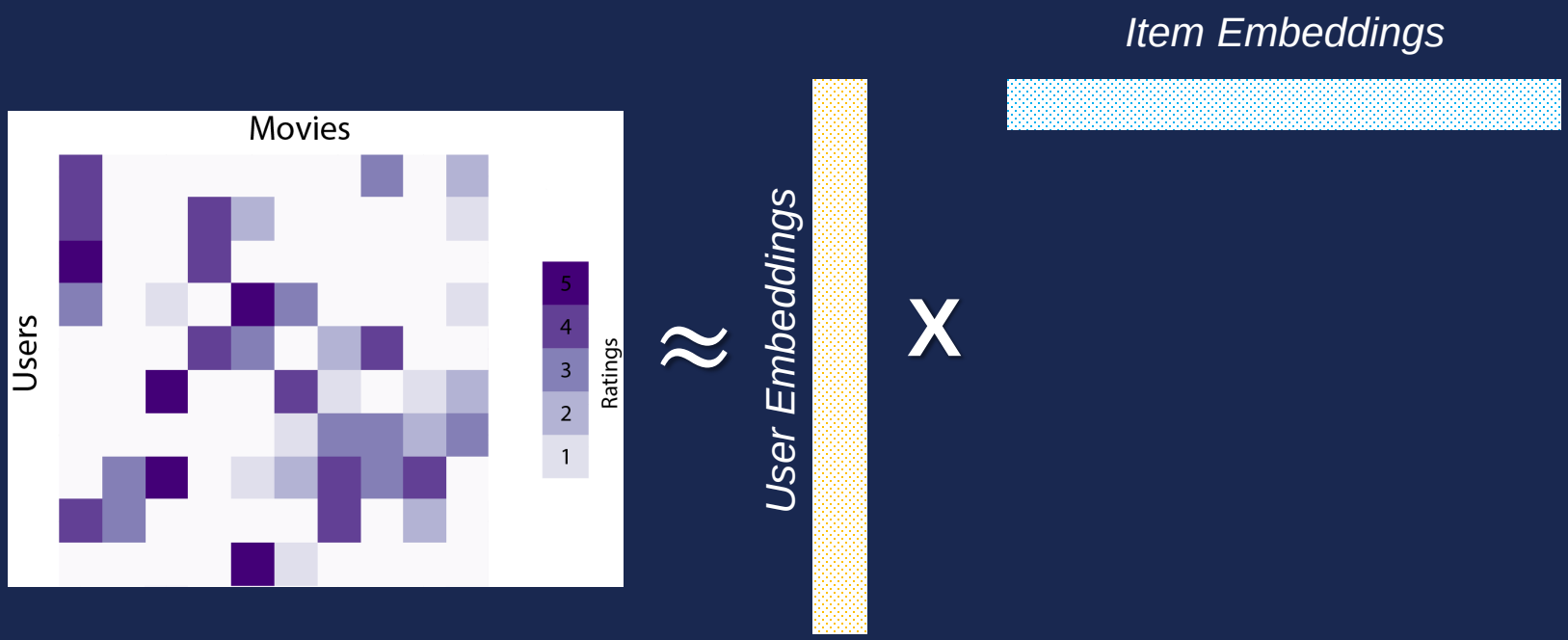
https://www.oreilly.com/ideas/deep-matrix-factorization-using-apache-mxnet?cmp=tw-data-na-article-engagement_sponsored+kibird

Naïve approach

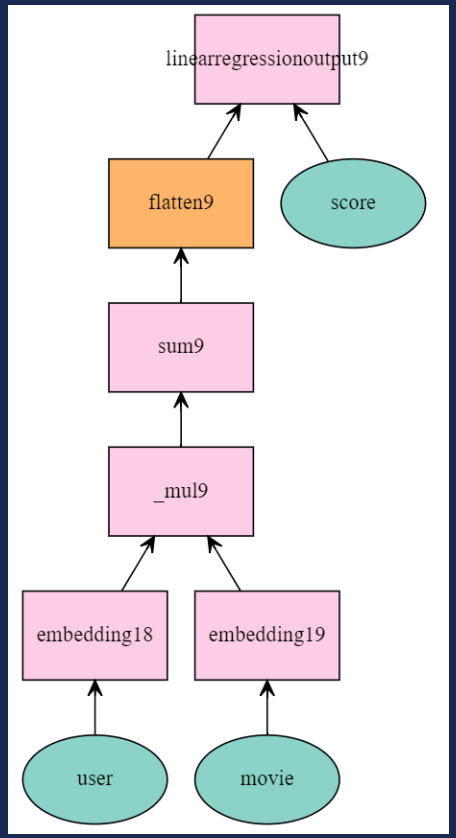
Linear model? [type of user, movie genre, etc.]

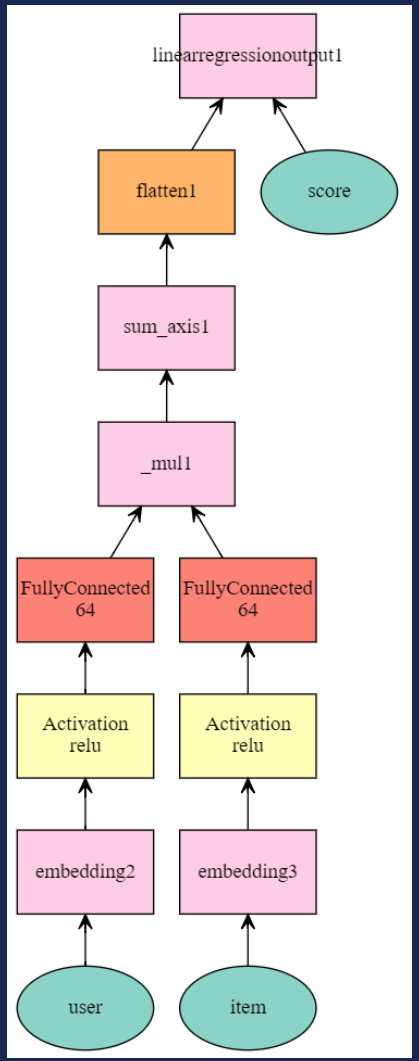
Polynomial model? [+interactions]

Matrix Factorization

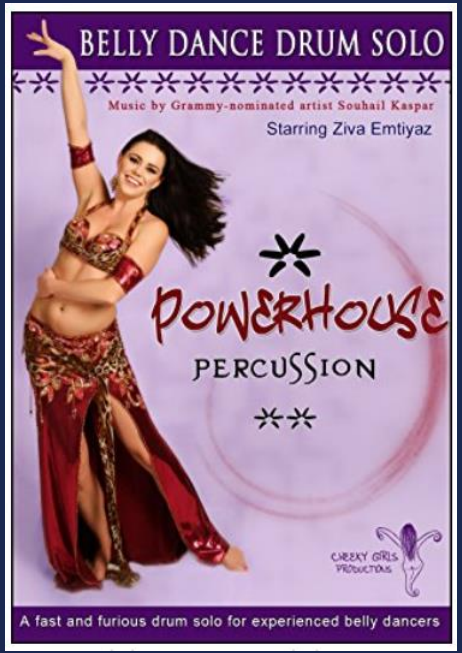


Matrix Factorization – “Neural Networks” Representation

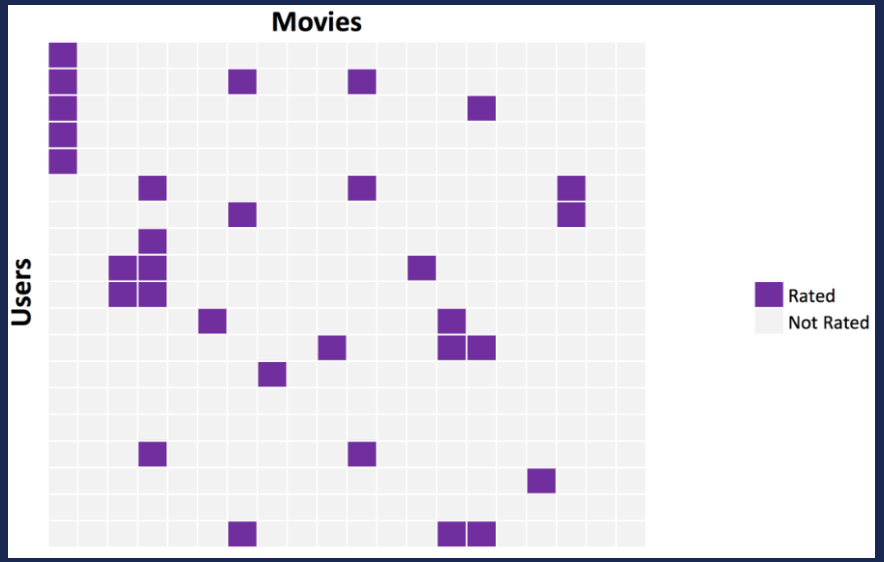




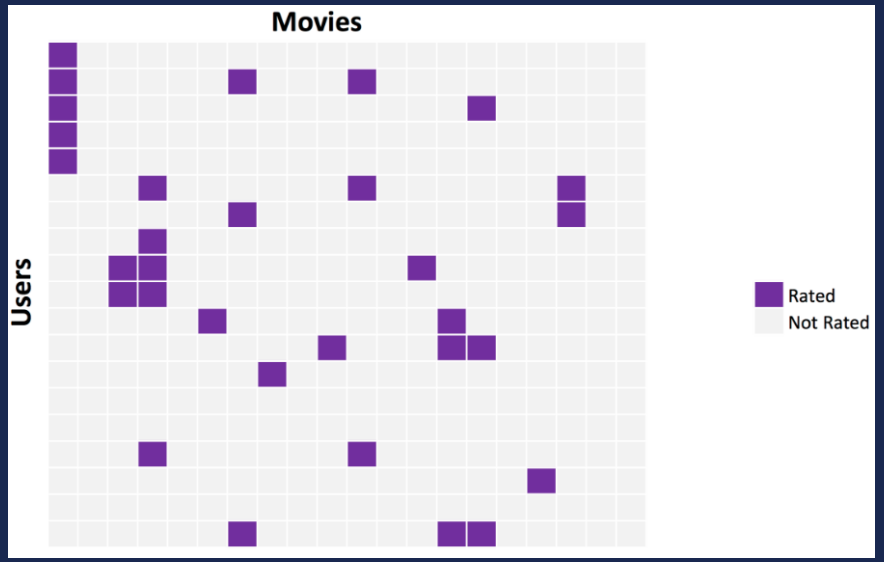
Deep Matrix Factorization

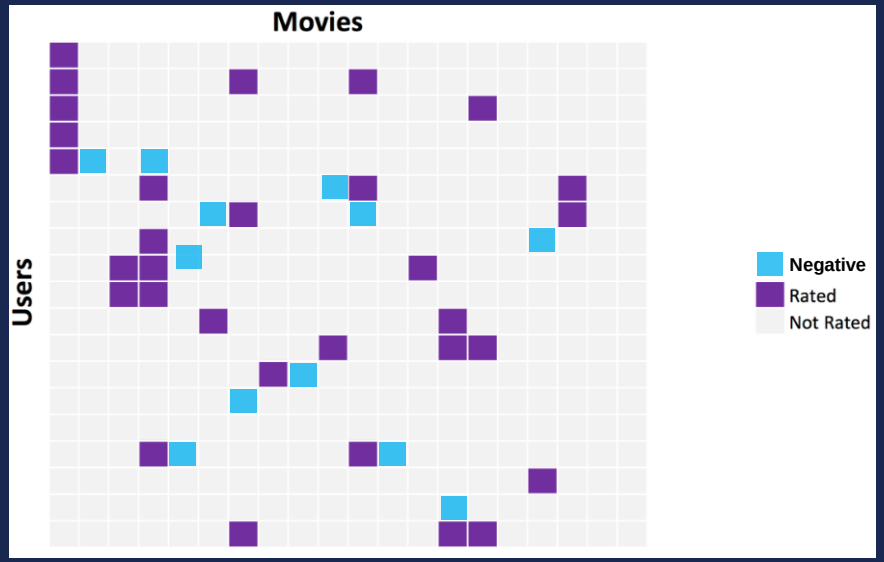


Binary Predictions



Binary Predictions

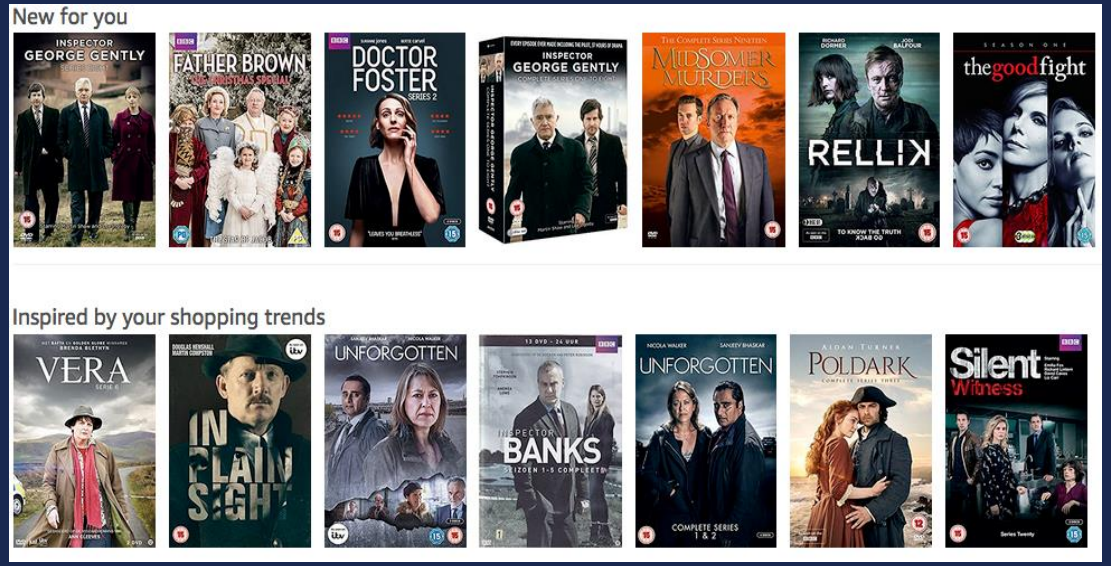




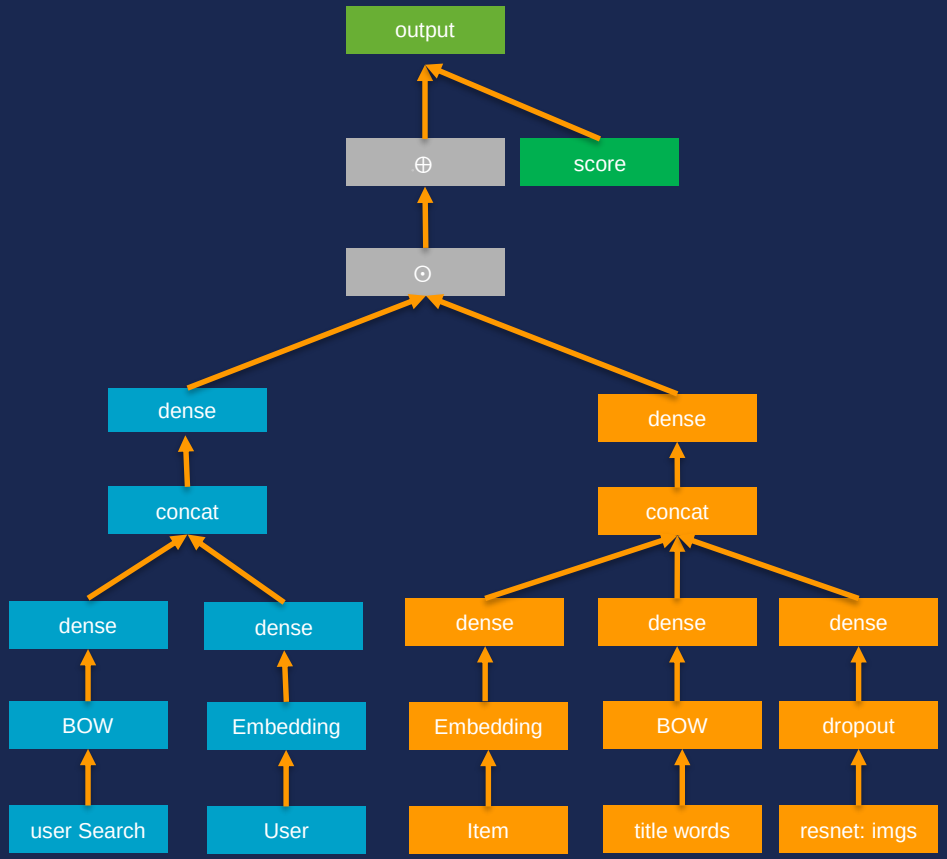
+Negative Sampling

Most of the Data is Still Untapped

- Images
- Titles
- Descriptions
- Reviews
- Episode Names



DSSM – Deep Structures Semantic Models



Which Technique to Choose? Roadmap Matrix

Iterative process	→	→	→	→
Data Available	Limited user data Binary user-item interaction	User data Additional user-item interaction	More user data Extensive item data	Extensive user data Extensive item data
Relevant Algorithms	Matrix Factorization Binary	Matrix Factorization Factorization Machines DiFacto	DSSM	Customized and more advanced DSSM
Relative Complexity	2	4	5	5
Deployment Considerations	• Historical data size – 30d / 60d / 1y… • Fine-tuning techniques (daily, weekly..) • Inference - compressed model? Tradeoff between model complexity and inference latency • Validation system setup • Iterate fast and simple			

Agenda

- Recommendation Engine – Why?
- Recommendation Engine – Common Techniques
- **Introducing Amazon SageMaker**
- Develop, Train & Deploy a Recommendation Engine in 15 minutes
- Customer Use Cases

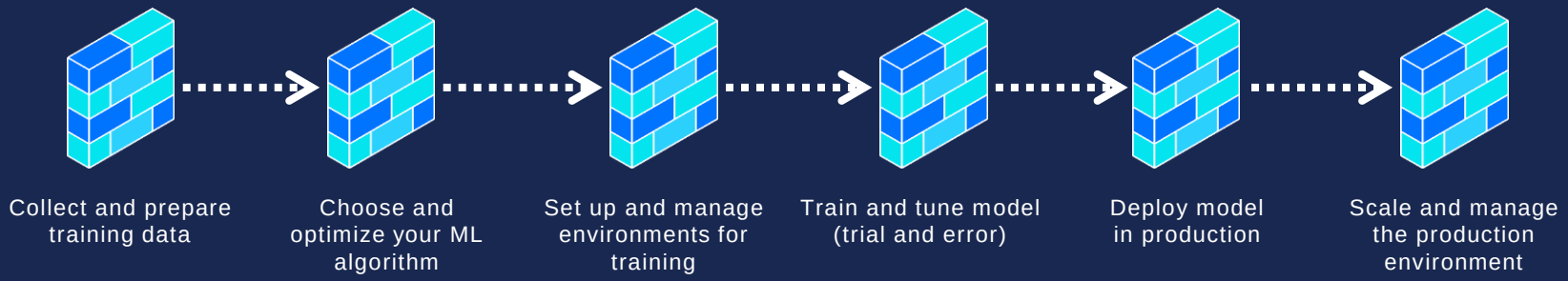
ML @ AWS: Our mission

Put machine learning in the hands of every developer and data scientist

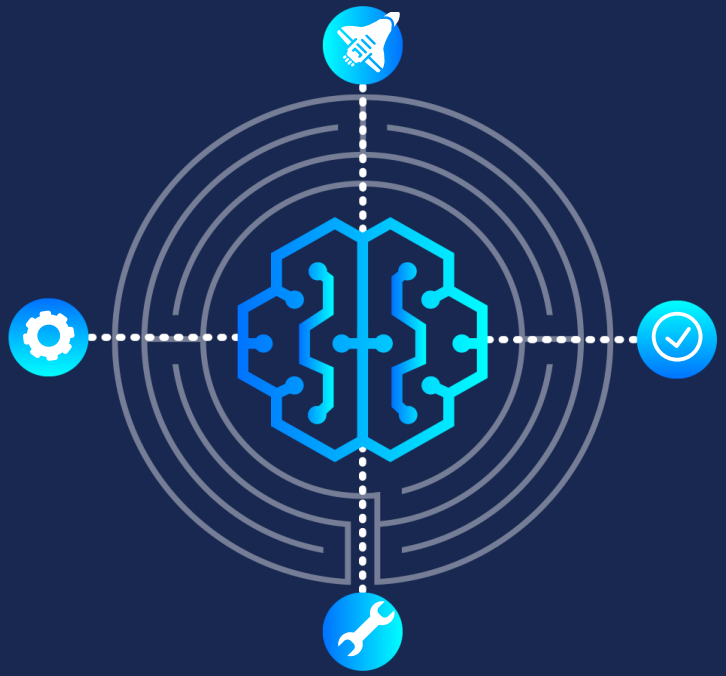
Customer Running ML on AWS Today

aws SUMMIT

ML is still too complicated for everyday developers and data scientists

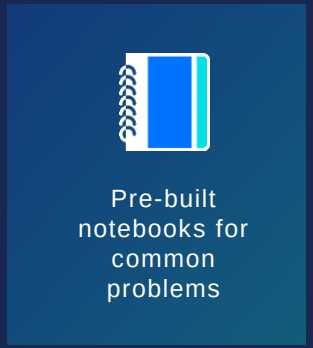


Amazon SageMaker



Easily build, train, and deploy machine learning models

Amazon SageMaker



BUILD

Choose and optimize your ML algorithm

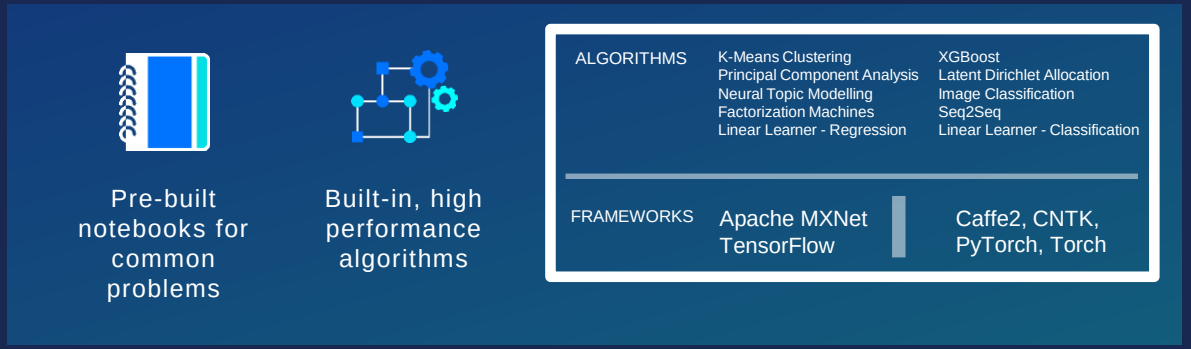
Set up and manage environments for training

Train and tune model (trial and error)

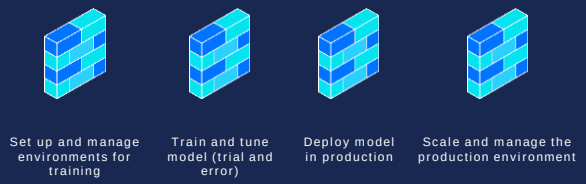
Deploy model in production

Scale and manage the production environment

Amazon SageMaker



BUILD



Amazon SageMaker

Pre-built notebooks for common problems

Built-in, high performance algorithms

One-click training

Train and tune model (trial and error)

Deploy model in production

Scale and manage the production environment

BUILD

TRAIN

Amazon SageMaker

Pre-built notebooks for common problems

Built-in, high performance algorithms

One-click training

Hyperparameter optimization

BUILD

TRAIN

Deploy model in production

Scale and manage the production environment

Amazon SageMaker

Pre-built notebooks for common problems

Built-in, high performance algorithms

One-click training

Hyperparameter optimization

One-click deployment

Scale and manage the production environment

BUILD

TRAIN

DEPLOY

Amazon SageMaker

BUILD

- Pre-built notebooks for common problems
- Built-in, high performance algorithms

TRAIN

- One-click training
- Hyperparameter optimization

DEPLOY

- One-click deployment
- Fully managed hosting with auto-scaling

Agenda

- Recommendation Engine – Why?
- Recommendation Engine – Common Techniques
- Introducing Amazon SageMaker
- **Develop, Train & Deploy a Recommendation Engine in 15 minutes**
- Customer Use Cases

console

Agenda

- Recommendation Engine – Why?
- Recommendation Engine – Common Techniques
- Introducing Amazon SageMaker
- Develop, Train & Deploy a Recommendation Engine in 15 minutes
- **Customer Use Cases**

Customers Use Cases

"Erento's in-house Data Science team is using Amazon SageMaker to build and deploy ML models to solve item availability and decrease the enquiry-to-offer time through a recommendation system, which suggests similar items that are available and increases the chance for a successful booking. Using Amazon SageMaker reduced our recommendation system building time from half a year to few weeks and reduced the algorithm training time from hours to few seconds. It also helped us reduce dependencies between projects, which has streamlined our whole pre-deployment process."

- Wassim Zoghlami, Data Scientist Engineer at Erento

"Using machine learning, we can provide better recommendations for our clients and enhance their customer experience. The AWS ML Acceleration Program delivered by the Professional Services Team, was really useful and suited our business needs. We believe that with Amazon SageMaker we can build a great recommendation system, and will be able to scale our ML training and deployment jobs in a more simple and faster way."

- Igor Veremchuk - Director of Engineering at Datajet

"Once we at HolidayPirates decided to take a strategic step towards personalization, we wanted to move fast. With the help of AWS Professional Services and the account team introducing us to Amazon SageMaker we are now able to develop, train and deploy recommendation system models in a very short time and independently from any other department. We no longer need to wear the hats of IT, big data, data science etc, and we can focus on what is important for our customers and enhance their user experience."

- Bojan Kostic, Data Team Lead at HolidayPirates

References

- https://www.oreilly.com/ideas/deep-matrix-factorization-using-apache-mxnet
- https://github.com/apache/incubator-mxnet
- https://github.com/awslabs/amazon-sagemaker-examples
- https://www.csie.ntu.edu.tw/~b97053/paper/Rendle2010FM.pdf
- https://www.youtube.com/watch?v=cftJAuwKWkA
- https://www.youtube.com/watch?v=1cRGpDXTJC8&t=640s

GO BUILD