Presented by:

ODSC

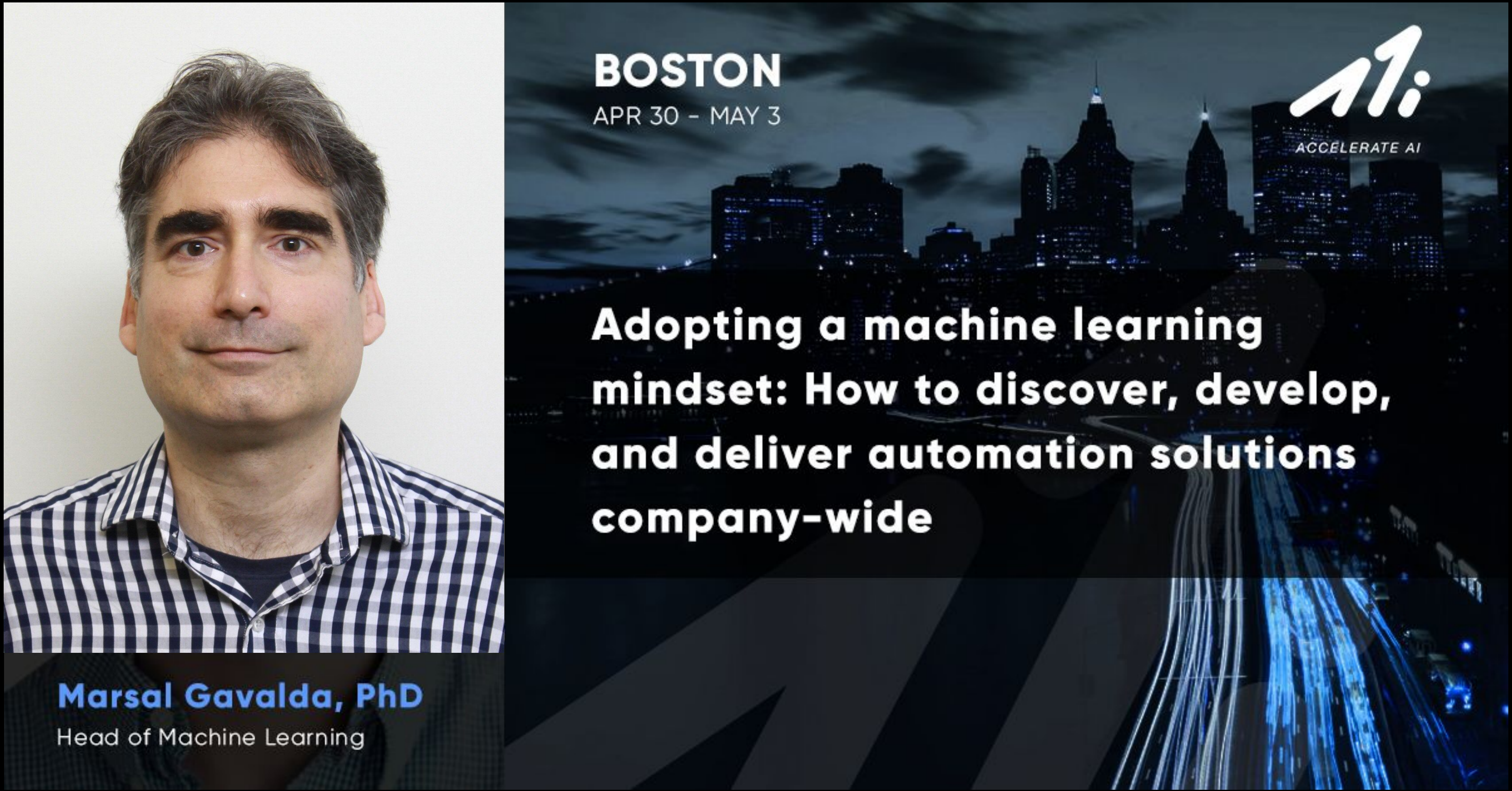
ACCELERATE AI

ACCELERATE AI EAST

BOSTON | April 30-May 4

2019

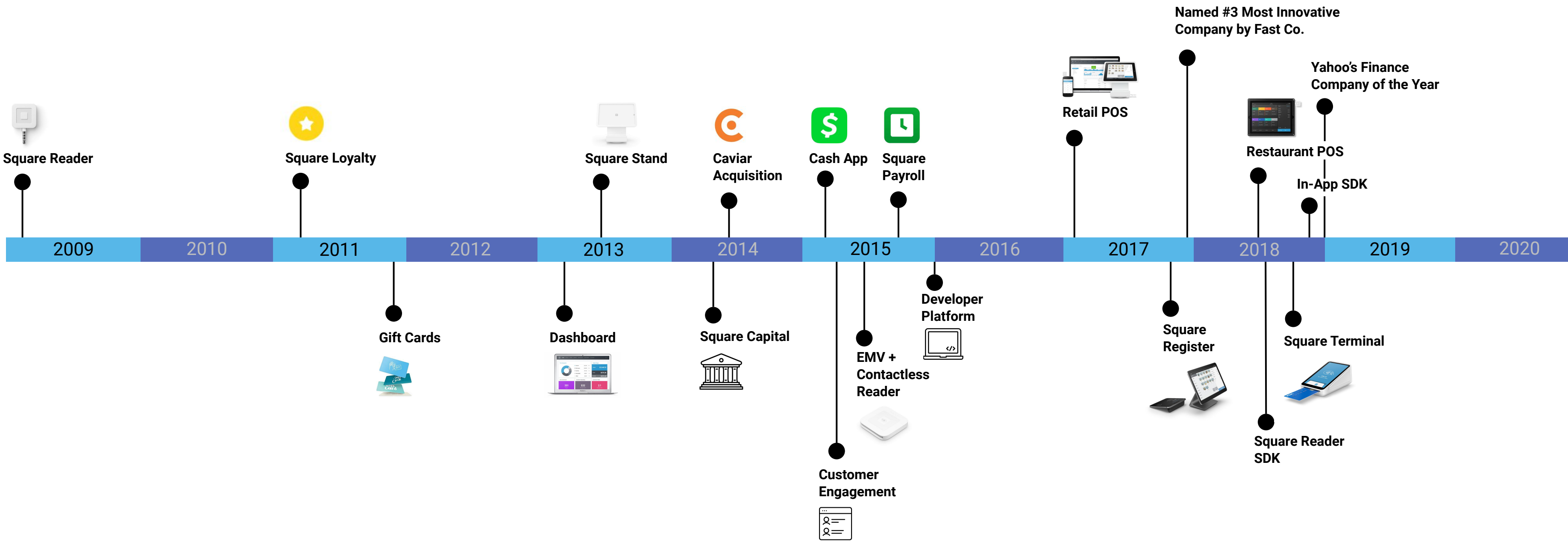
THE LEADING DATA SCIENCE CONFERENCE


BOSTON
APR 30 - MAY 3
ACCELERATE AI
Adopting a machine learning mindset: How to discover, develop, and deliver automation solutions company-wide
Marsal Gavalda, PhD
Head of Machine Learning


$15.52
Insert, Tap or Swipe
AMERICAN EXPRESS
Pay with Touch ID


RAISIN OATMEAL
CHRISTMAS PUDDING
Tips

Square Timeline



Payments

Process payments anywhere you sell.

Learn More ›

Deposits

Get your money the next business day.

Learn More ›

Invoices

Send invoices free and get paid fast.

Learn More ›

Gift Cards

Sell customized gift cards and boost sales.

Learn More ›

Point of Sale

Run your entire point of sale from the free Square app.

Learn More ›

Employee Management

Set permissions and keep track of your team.

Learn More ›

Location Management

Manage multiple locations from one account.

Learn More ›

Analytics

Get real-time sales data and customer insights.

Learn More ›

Dashboard

Manage your entire business from one place.

Learn More ›

Dashboard App

Keep your business in the palm of your hand.

Learn More ›

Square Marketing

Send customized, targeted email.

Learn More ›

Loyalty

Reward repeat customers with a loyalty program.

Learn More ›

Payroll

Run payroll in just a few clicks.

Learn More ›

Capital

Get a loan to grow your business.

Learn More ›

App Marketplace

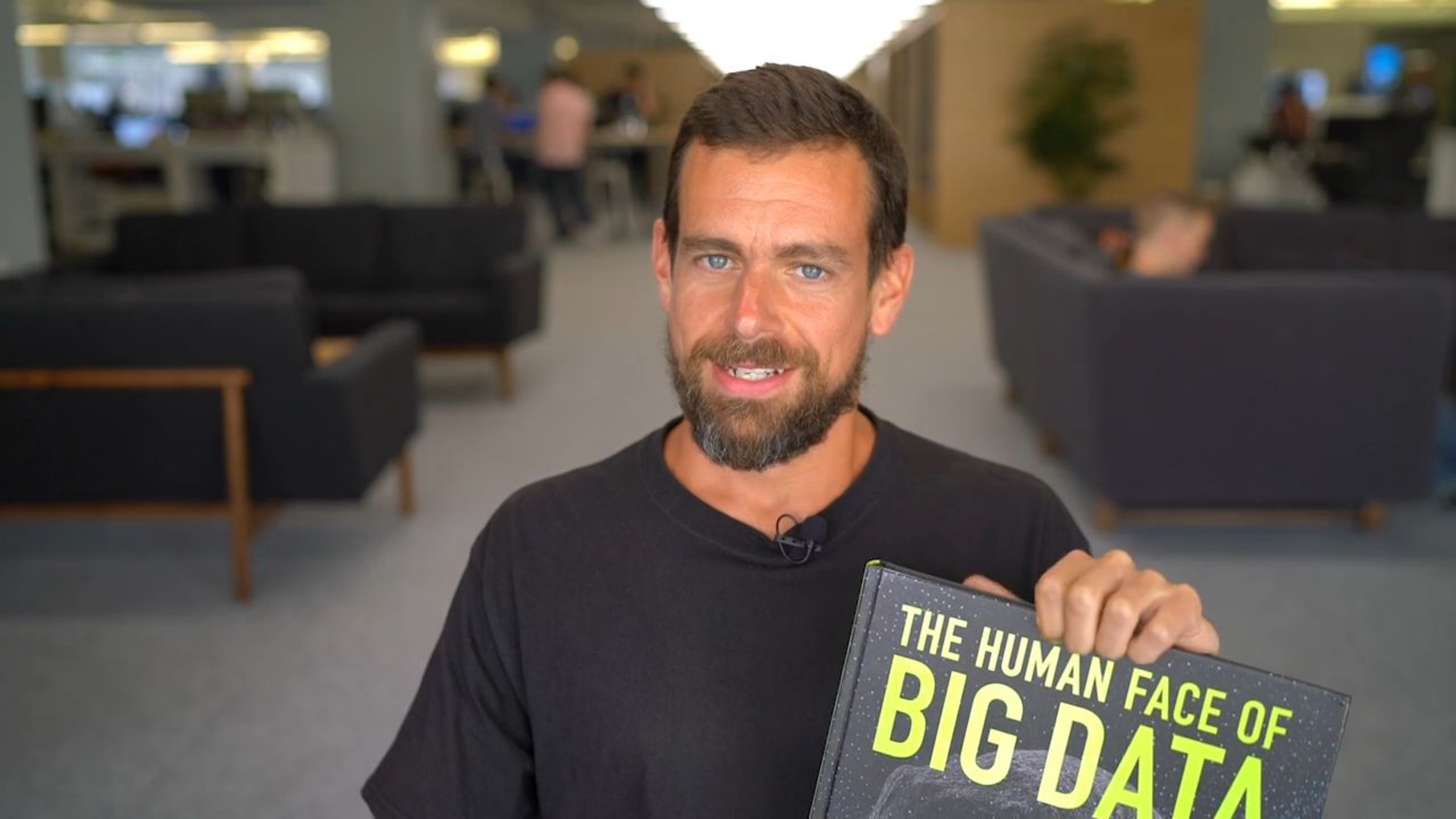
Seamlessly connect apps to your point of sale.

Learn More ›

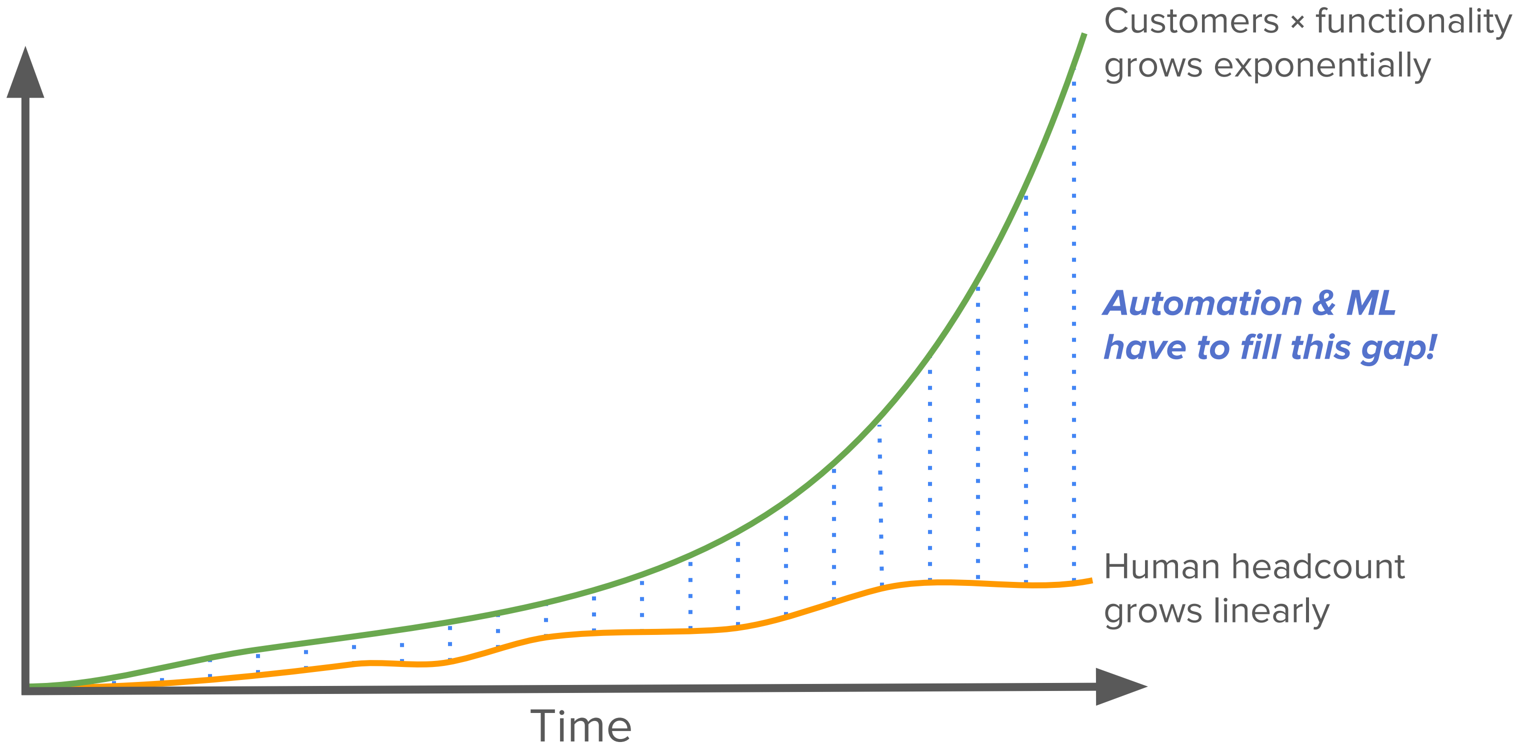
Developer APIs

Create a custom solution with Square's API.

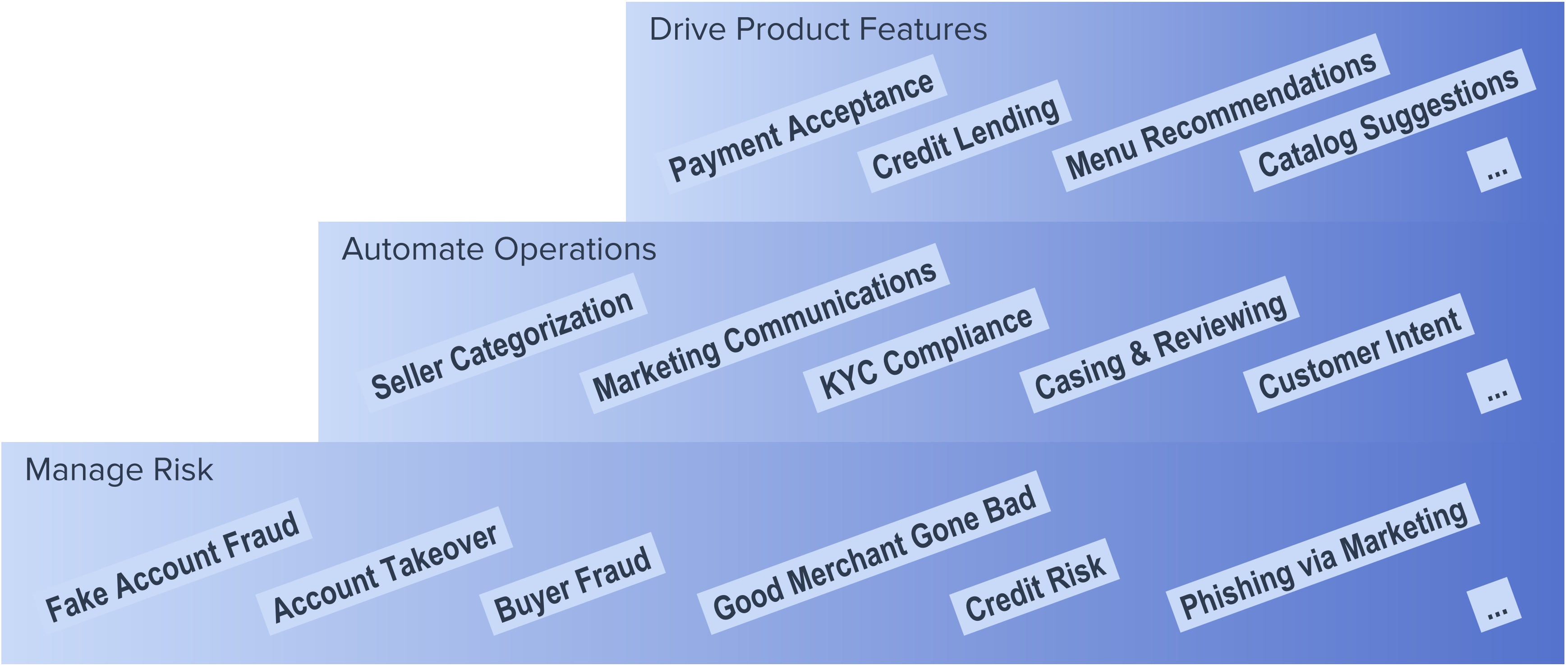
Learn More ›


THE HUMAN FACE OF
BIG DATA

Automation/ML drives productivity



Machine Learning at Square

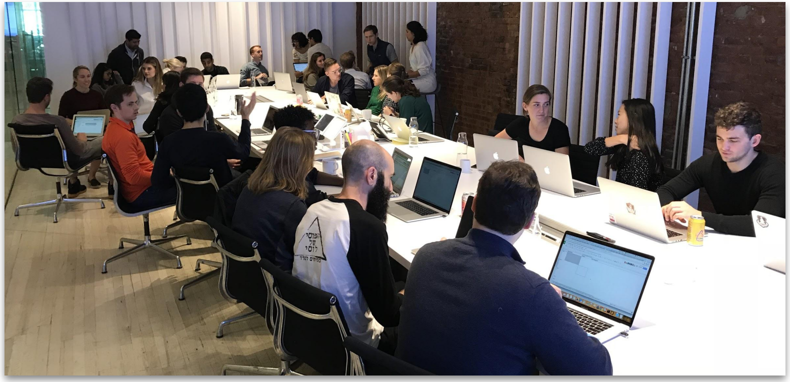


Adopting a machine learning mindset

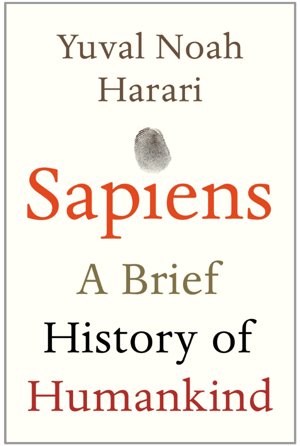
ML for Everyone

"Machine Learning for Everyone"

- The nuts and bolts of machine learning
- How we use ML at Square
- Your role (even if you are not an engineer or data scientist!) in bringing ML to life at Square
- Brainstorming on possibilities to apply ML to your day to day



Key phases in human history



Cognitive revolution [200k years ago]

- Evolution of *Homo sapiens*: abstract thinking, long-term planning
- Hunter-gatherer societies expand throughout the world

Agricultural revolution [12k years ago]

- Stable food source dramatically increases tribe size, density, stratification.
- Invention and adoption of ***writing*** and ***money*** leads to cities, kingdoms, political alliances and commercial links

Scientific revolution [400 years ago]

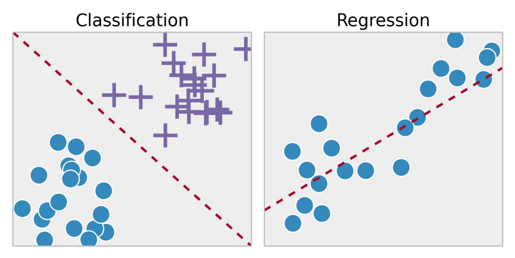
- Scientific method: $\textit{Knowledge} = \textit{Empirical Data} \times \textit{Mathematics}$
- Leads to modern medicine, industrial revolution, information technology, and... first glimpses of AI

Machine learning algorithms

Supervised Learning

Algorithm learns from labeled examples

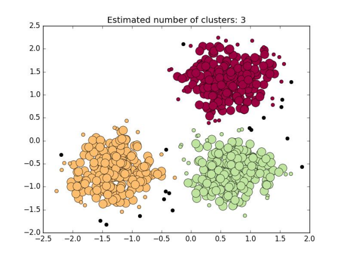
- *Classification*
  - E.g., spam email
- *Regression*
  - E.g., home price



Unsupervised Learning

Algorithm identifies patterns in the data

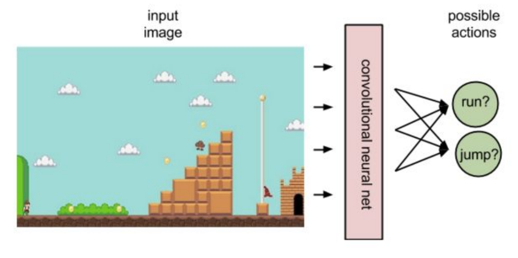
- *Clustering*
  - E.g., topics in Support emails
- *Anomaly detection*
  - E.g., fraudulent transaction
- *Recommender system*
  - E.g., personalized suggestions



Reinforcement Learning

Algorithm learns optimal behavior over longer-time cumulative reward

- *Robotics*
- *Game playing*

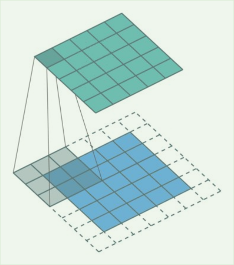


How ML models are built

Training

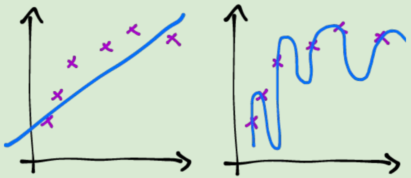
Raw Data →

Feature Extraction
- Map raw data to useful signals
- E.g, aggregate into averages, or cluster into similar groups



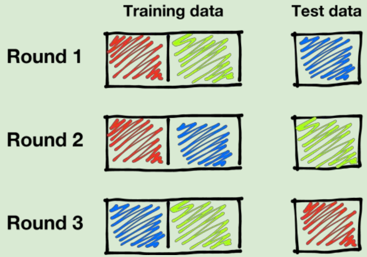
→

Regularization
- Optimize model complexity, bias, variance
- Avoid underfitting or overfitting training data



→

Cross-validation
- Partition dataset into train / test subsets
- Evaluate on test
- Repeat multiple times



→ **ML Model**

Inference

New datapoint → Features → ML Model → Prediction

Modeling exercise — **Estimate home price**

Model: How can we combine the features into a pricing model?

```
price =
  c1·v1 + … + cn·vn
```

$$price = c_1 \cdot v_1 + \ldots + c_n \cdot v_n$$

v_i: Values of features

c_i: Coefficients that model

- positive vs. negative effect feature has on the price
- scaling factor

Accuracy: How do we measure the accuracy of a model?

$$error = |y_1 - \hat{y}_1| + \ldots + |y_n - \hat{y}_n|$$

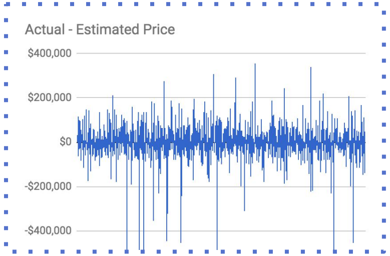
y_i: Actual price

$\hat{y}_i$: Estimated price

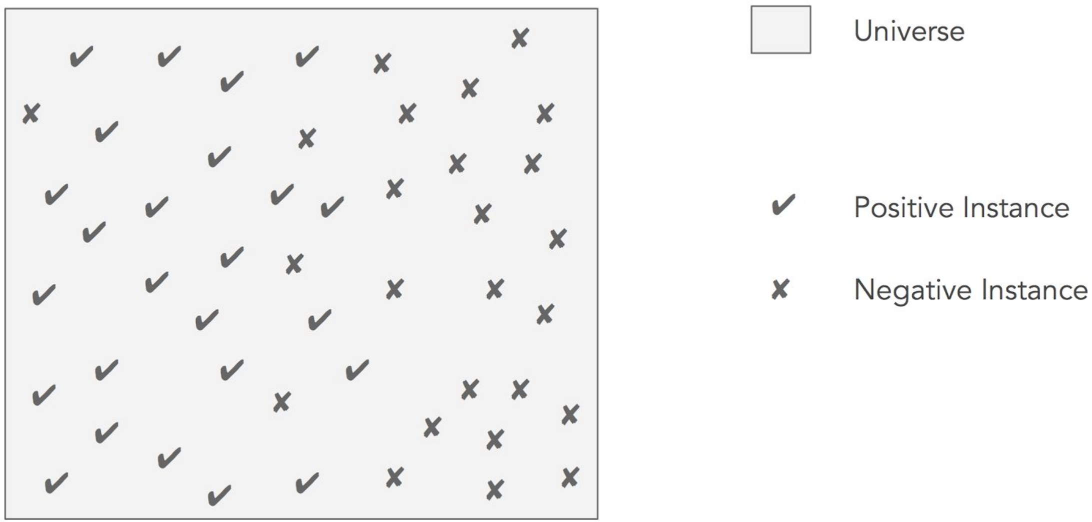
Take the average of the absolute value of the differences between actual and estimated prices.

Training: How do we learn from the data?

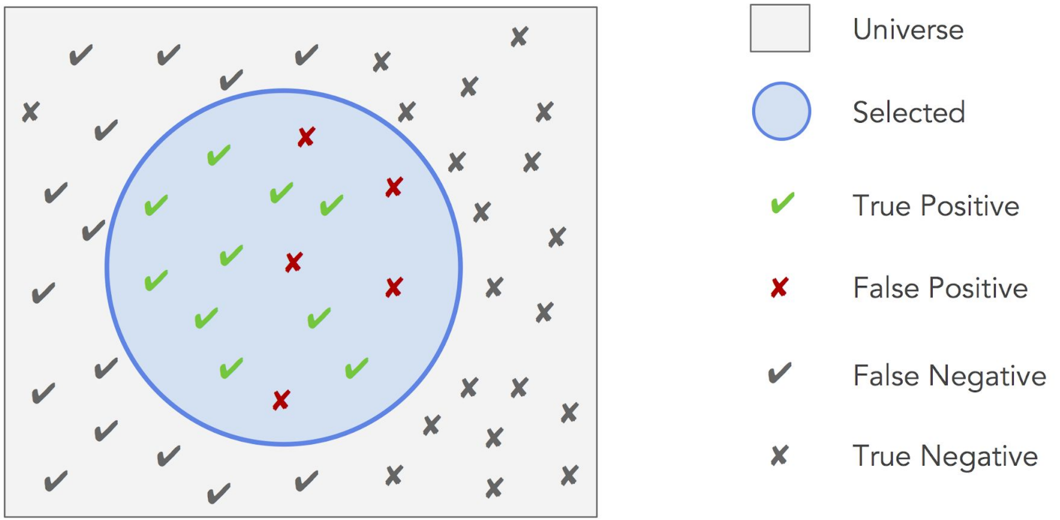
Adjust model coefficients to minimize error.



Measuring accuracy



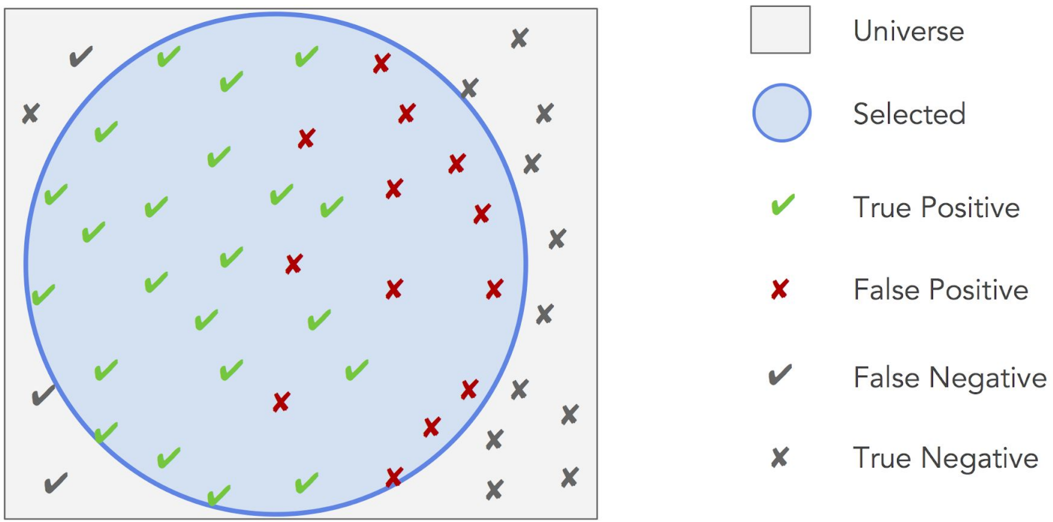
Accuracy: **Precision vs. Recall**



Accuracy

- **Precision**: TP / (TP+FP) = 10/(10+5) = 10/15 = **66%**
- **Recall**: TP / (TP+FN) = 10/(10+15) = 10/25 = **40%**

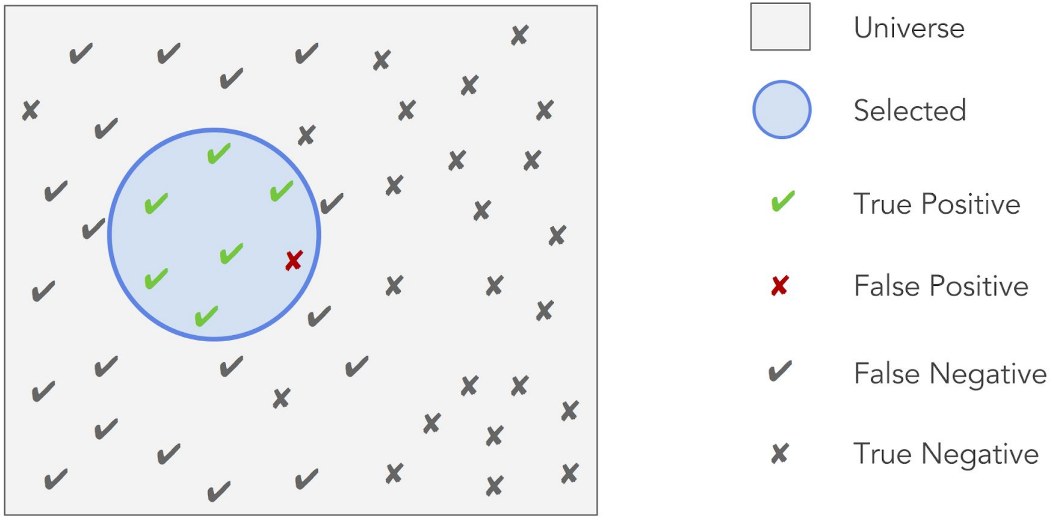
Low-confidence threshold: **Precision↓, Recall↑**



Accuracy

- **Precision**: `TP / (TP+FP) = 22/(22+13) = 22/35 = ` **63%**
- **Recall**: `TP / (TP+FN) = 22/(22+3) = 22/25 = ` **88%**

High-confidence threshold: **Precision↑, Recall↓**



Accuracy

- **Precision**: `TP / (TP+FP) = 6/(6+1) = 6/7 = ` **86%**
- **Recall**: `TP / (TP+FN) = 6/(6+19) = 6/25 = ` **24%**

Which model is best?

To select the "best" model we need to weigh the *cost* of a

- **False Positive** (delinquent loan): liable for entire loan amount ➔ ***Let's favor precision!*** *vs.*
- **False Negative** (good seller rejected): lost profit, bad UX ➔ ***Let's favor recall!***

Ultimately, it's a **data-informed business decision**.

Also important:

- No demographic bias
- Model interpretability

ML Mindset — **Identifying automatable decisions**

1. **Define customer problem**
What are we trying to solve? Has someone already done it? Is it worth solving? Is ML machine learning appropriate?

2. **Collect data**
If we don't yet have the relevant data, work with engineering to implement the necessary instrumentation and collection.

3. **Build model**
We use features derived from the data to train a machine learning model. The model is evaluated on accuracy, runtime speed, and interpretability.

4. **Iterate**
We improve the model by training with more data, tweaking the features, etc.

ML Mindset — Identifying automatable decisions

Square's Operating Principles: **We ship remarkable solutions by**

1. **Define customer problem**
What are we trying to solve? Has someone already done it? Is it ... solving? Is machi... appro...

1. Understanding someone's struggle

3. **Build model**
We use features derived from the data to train a machine le... model. The model is ... accuracy, ...peed, and interpretability.

3. Taking principled risks

2. **Collect data**
If we don't yet have the relevant data, work with engineering to imp... it the necessary ins... and collec...

2. Seeking divergent perspectives

4. **Iterate**
We improve the model by training with more data, tweaking th... atures, etc.

4. Inventing and learning

ML Mindset — **Your role as a non-tech Square**

1. **Define customer problem**
Think about the needs of your customers (both internal and external). Apply JTBD framework to reason about these needs. *How will ML help?*

2. **Collect data**
As a Subject Matter Expert, help identify key pieces of information that are relevant to model the task at hand. *Typically there are hundreds of potential features, but only a handful prove useful for a particular task.*

3. **Build model**
Provide context and guidance to the data scientist building the model. Help interpret the initial results.

4. **Iterate**
Keep an eye on model performance in the wild. Look at false positives and false negatives to see if there are any patterns we can use to correct the model and improve accuracy over time.

Adopting a machine learning mindset

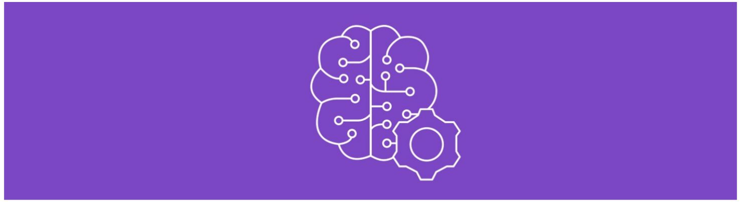
ML Technical Bootcamp

ML Technical Bootcamp

Video lectures, presentations, quizzes, and coding exercises on:

- DS and ML at Square
- Experimental Design and Pandas
- Statistics Fundamentals
- Regression and Analysis
- Evaluating Model Fit
- Classification
- Logistic Regression
- Communicating Results
- Clustering
- Decision Trees and Random Forests
- NLP and Text Classification
- Latent Variables and NLP
- Time Series Analysis and Modeling
- ML and Engineering
- Neural Networks and Deep Learning

Machine Learning Bootcamp



What is it?

This four-week, 12-class course starts with a high level overview of the ML and DS field, followed by a deeper introduction to fundamentals of data science and ML theory. We'll do a deep dive into commonly used techniques expecting hands-on homework by the participants. All presented techniques will be explored using practical examples and applications to everyday work at Square. The last week is a hack week in which participants work on their own projects, intended to motivate further exploration after the course.

What to Expect?

There are currently five modules available for online learning. Each module has quizzes and coding exercises.

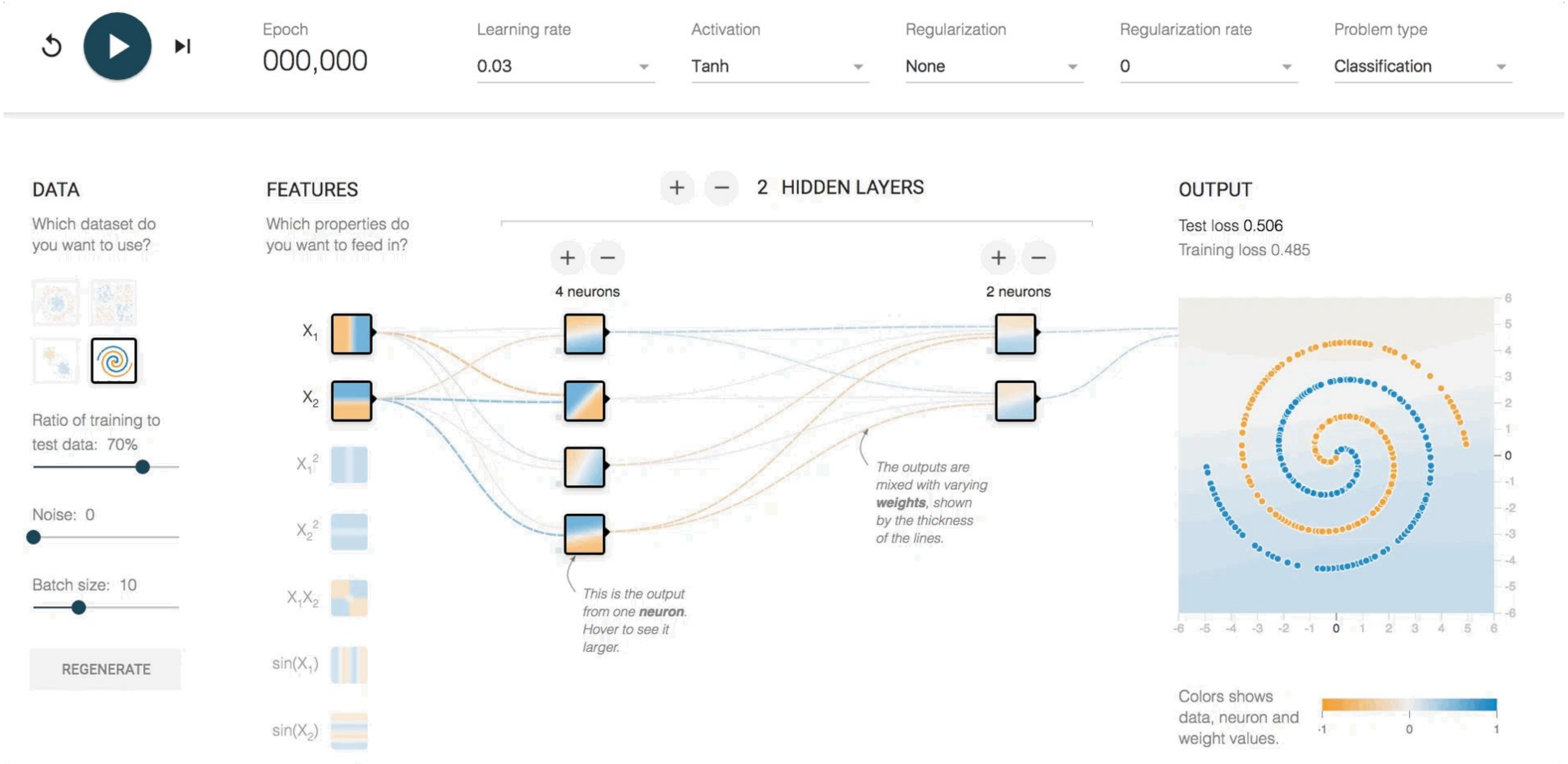
How Do I Enroll and Access Content?

- Click on the Module links listed below to deep link directly to each module in Workday.
- Having difficulty? Here are instructions on how to enroll in these workshops in Workday Learning.

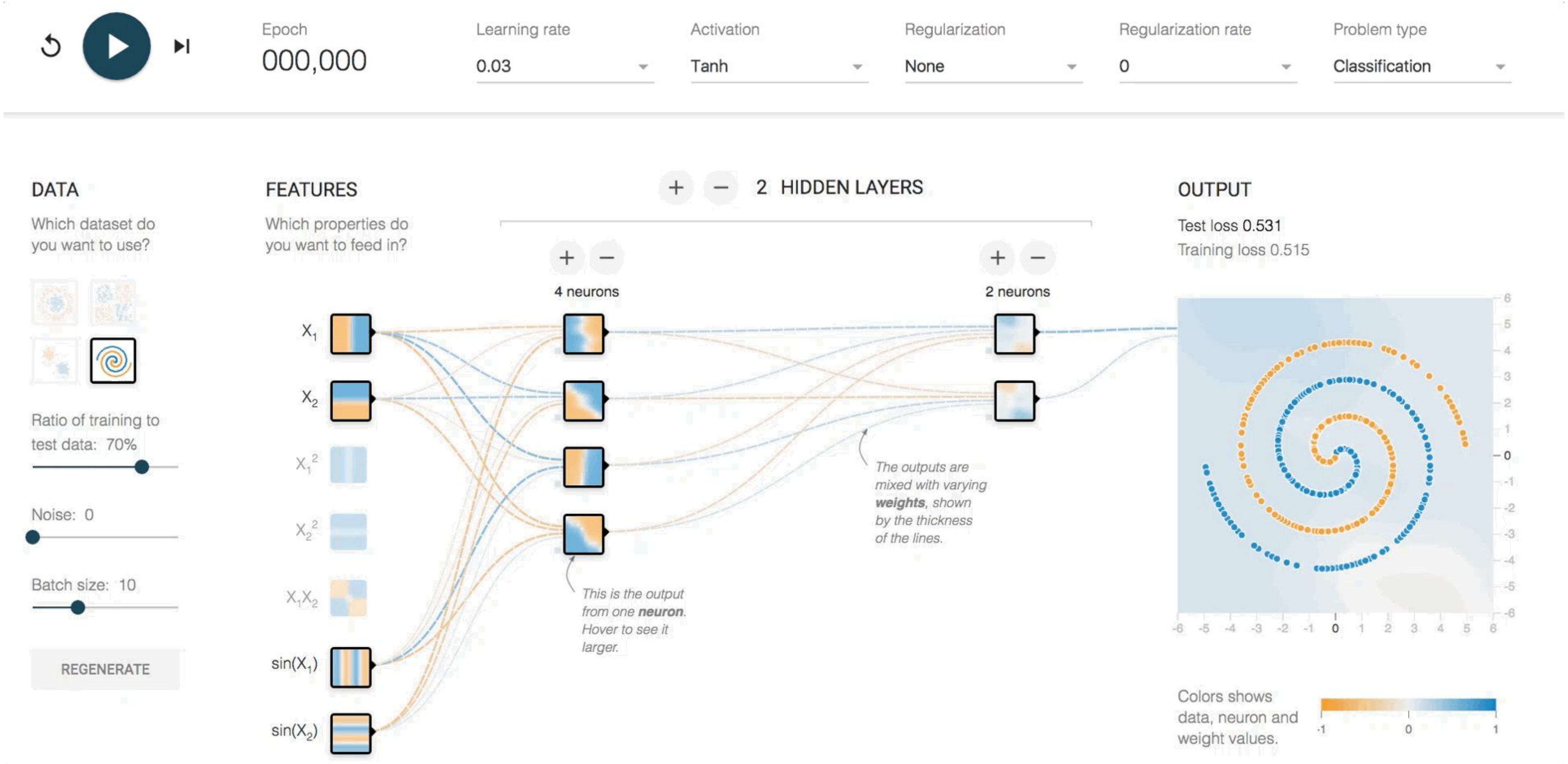
Objectives

- Allow a deeper exploration of content presented in the one-day workshop.
- Gain a better understanding to common techniques used in DS and ML.
- Understand how these techniques are used at Square.
- Provide practical experience with tools.

Training a neural network — **Unable to learn!**



Training a neural network — **Adding features**



Adopting a machine learning mindset

ML for Designers and PMs

ML-driven product — **The wow square**

Design
Will users care?

Functionality
Aesthetics
Incorrect prediction behavior

Modeling
Can it be automated?

Data sources
Features & models
Accuracy

Engineering
Can it be built?

Robustness
Latency & throughput
Scalability

Analytics
Is it successful?

Instrumentation
Adoption metrics
Feedback for ML training

ML-driven product — **The wow square**

Design
Will users care?

Functionality
Aesthetics
Incorrect prediction behavior

Modeling
Can it be automated?

Data sources
Features & models
Accuracy

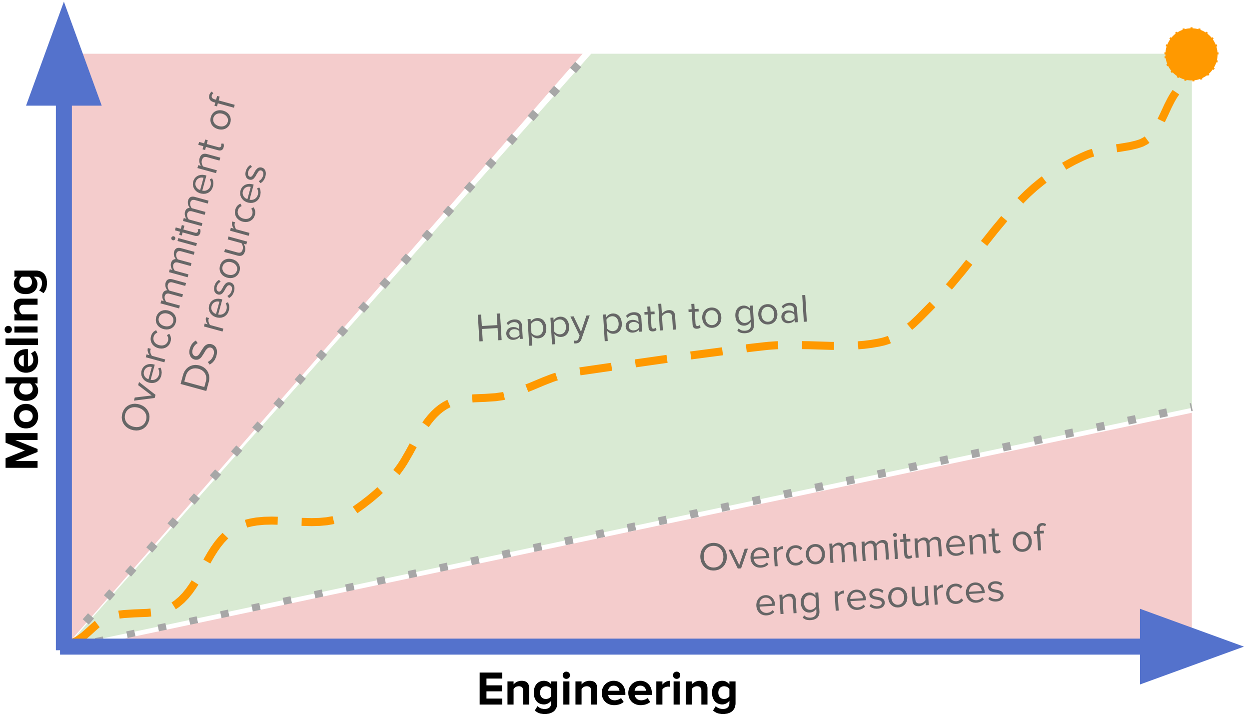
Engineering
Can it be built?

Robustness
Latency & throughput
Scalability

Analytics
Is it successful?

Instrumentation
Adoption metrics
Feedback for ML training

PMing ML project — **Modeling adds a new dimension!**



Increasing adoption of ML

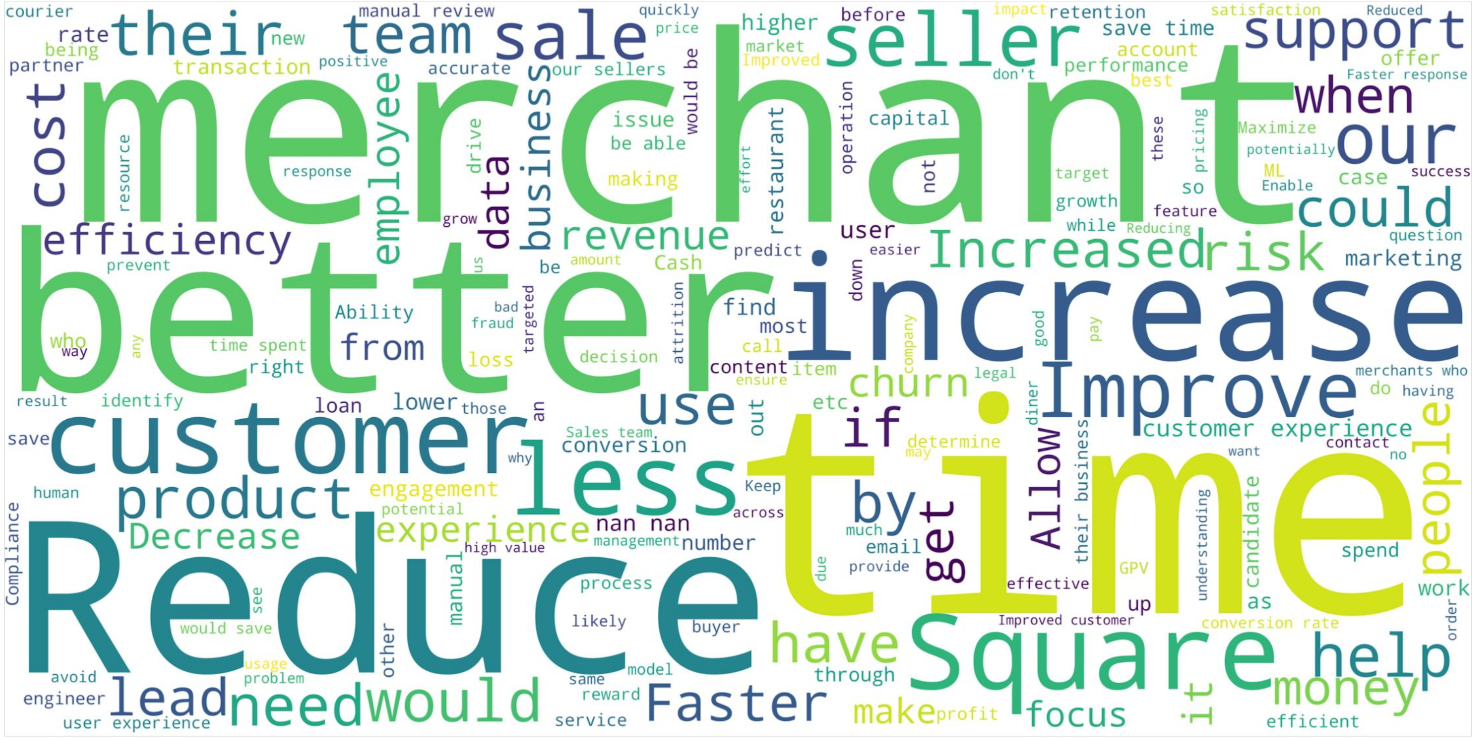
Brainstormed automation ideas

Brainstorming Time!

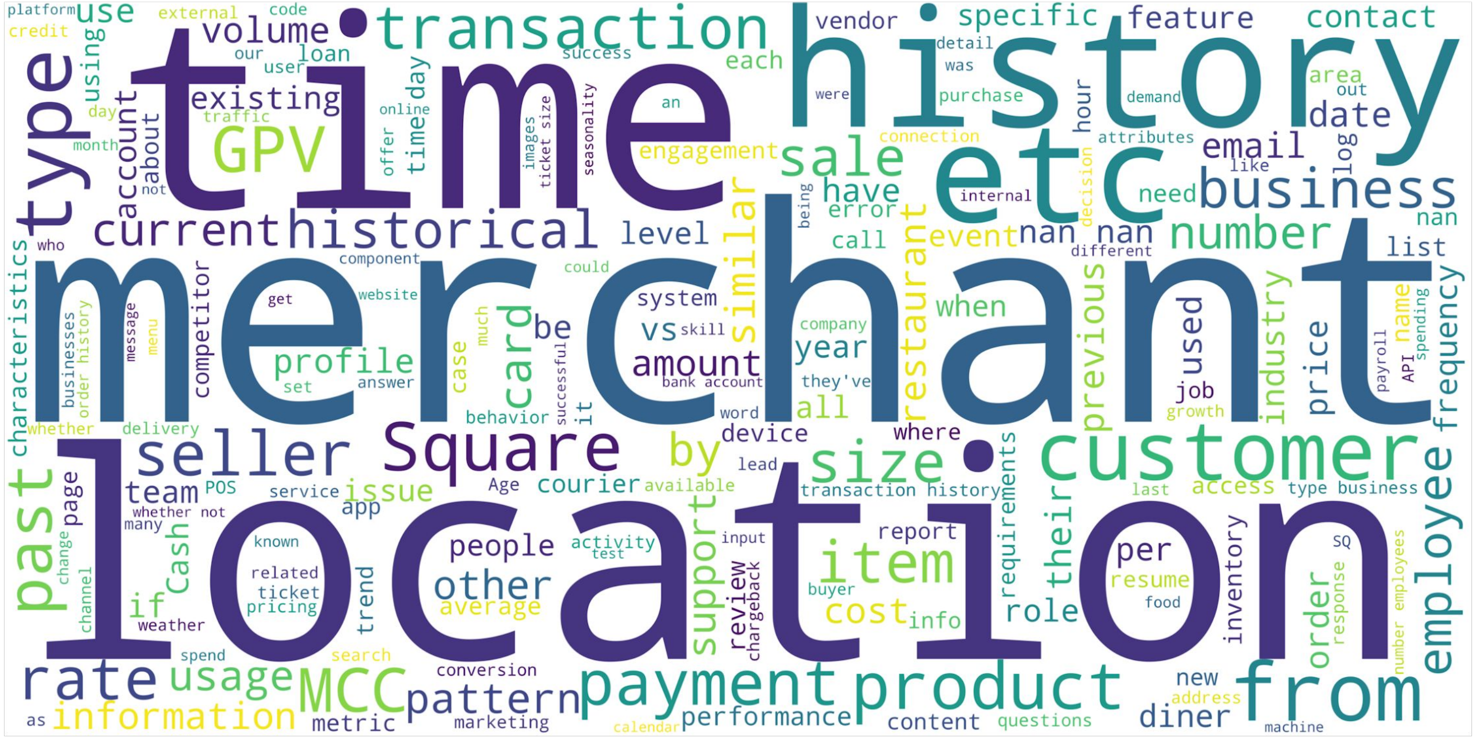
“What could ML do for my team or my customers?”

Automation ideas — **Description**

Automation ideas — **Impact**



Automation ideas — **Data**



Automation ideas — **Sample Themes**

Engineering processes

Automate *UI & browser testing*, *build process*, and *accessibility auditing.*

Customer success

Improve Support by *anticipating problems* a Seller may face and *automating responses.*

Internal operations

Improve the efficiency of internal processes such as *ranking sales leads* by conversion probability, automating the *scheduling of meeting rooms* and *candidate interviews*, and even optimizing *seating arrangements.*

Hardware

Automation ideas for hardware centered in forecasting and triaging failures, such as *proactively sending replacement units* and *reducing tamper false alarms.*

Automation ideas — **Technologies**

In terms of the *implementation* of the automation ideas, the most frequently occurring techniques are:

Recommender systems

E.g., suggest items that complement what's already in the shopping cart.

Natural language processing

E.g., automatically parse and respond to Support emails.

Time series forecasting

E.g., predict a Sellers' future sales to help with financial planning.

Anomaly detection

E.g., identify when a merchant's reader has failed.

Automation ideas — **Implementation**

Assemble DS + eng + PM action squads by drawing resources from most impacted teams

Prioritize

- Effort to prototype, productize
- Expected accuracy
- Impact on revenue, CSAT
- Impact on retention/growth

Implement

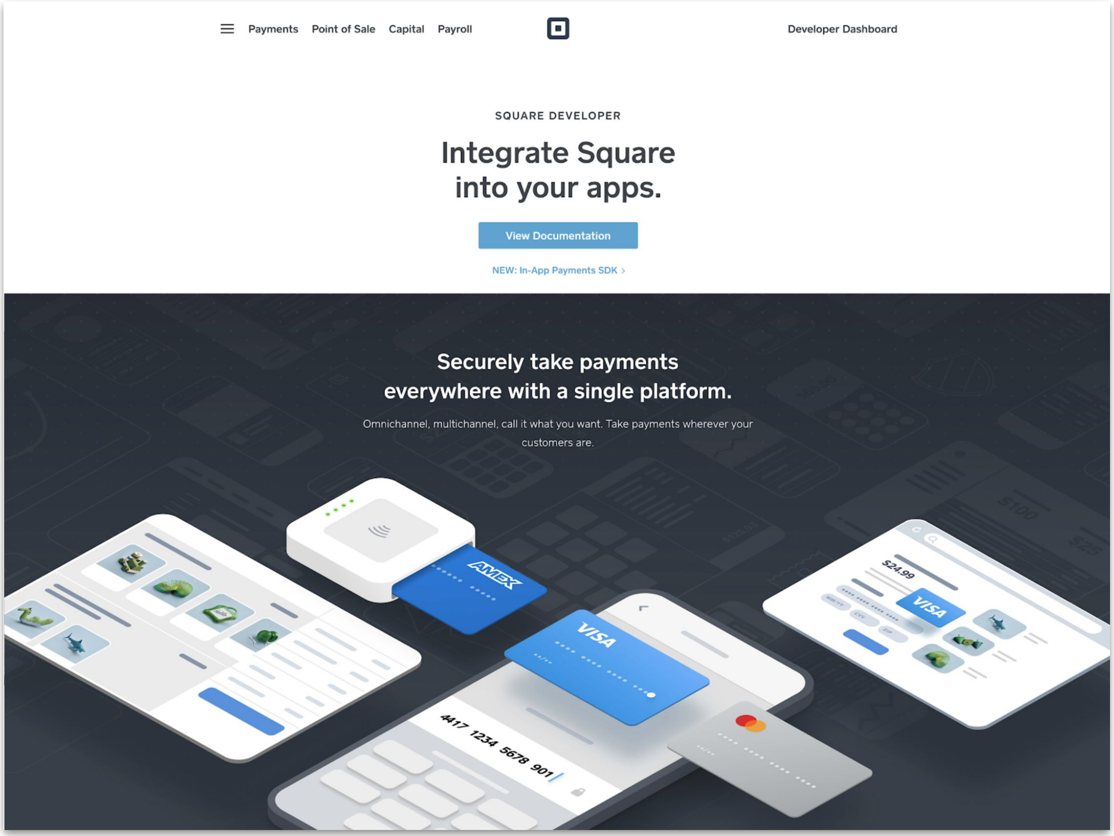
- Prototype idea (Wizard of Oz)
- Validate feasibility (train model)
- Quantify benefit
- Productize

Automation ideas — **API Scripting for Professional Services**

Problem
Square Professional Services performs manual tasks

Solution
Configurable script calls backend APIs

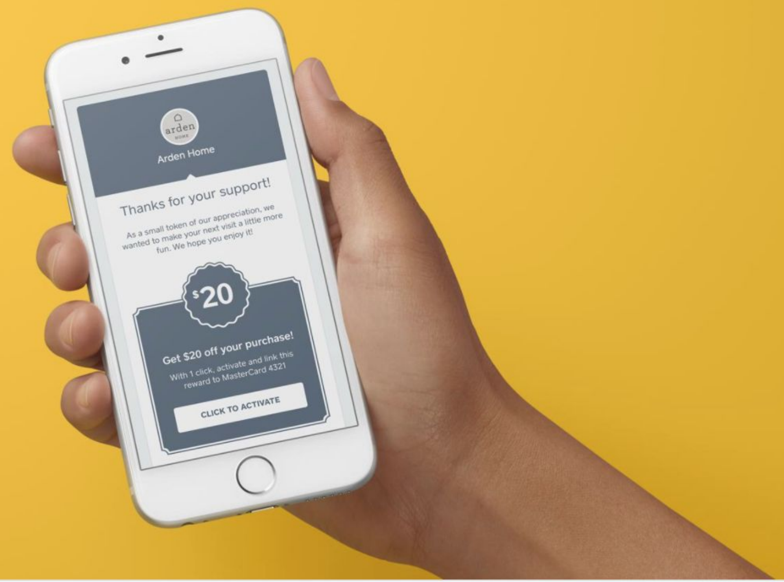
Benefits
Saves time and increases self-serve skills



SQUARE MARKETING

Grow your business with email marketing.

GET STARTED



OVERVIEW FEATURES TESTIMONIALS PRICING GET STARTED

Promote your company

Square's email marketing software helps you send email campaigns to keep your business top of mind. And now it can send automated campaigns for you.

Build your Customer Directory

Create new customer profiles at the point of sale to build your directory. Square also automatically adds key customer information with a swipe, dip, or tap.

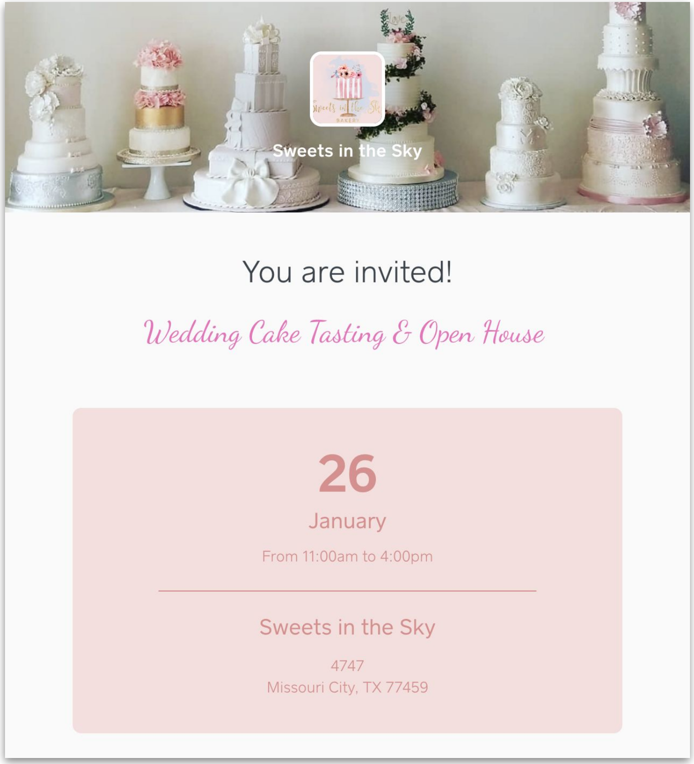
Send in minutes

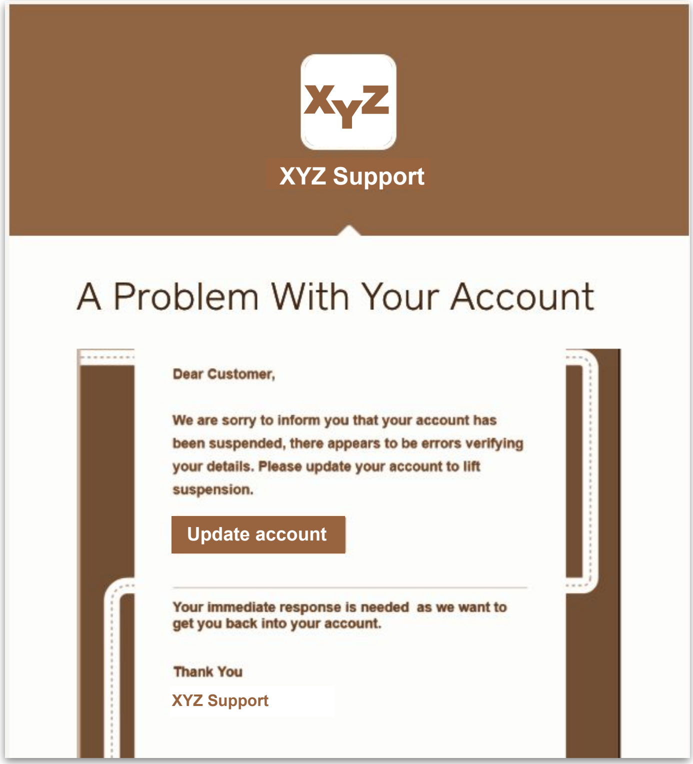
Customizable email templates and a distribution list built just for your business let you go from idea to sent email campaign in minutes.

Track results

Our email marketing tools show your return on investment for each email right in your Square Dashboard.

Automation ideas — **Phishing Detector**





0.06 → no phishing

0.98 → phishing!

Adopting a Machine Learning Mindset

Provide broad
ML training

Brainstorm
ML/automation ideas

Treat data as
first-class citizen

Weave ML into
products and processes

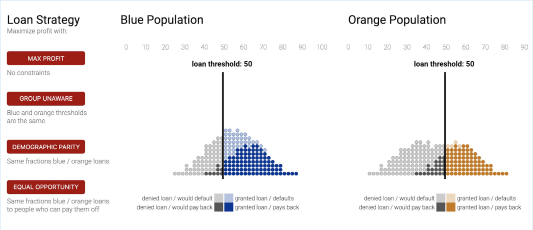
Follow ethical principles & practices

AI/ML Principles and Practices

The automated decisioning systems we build should be...

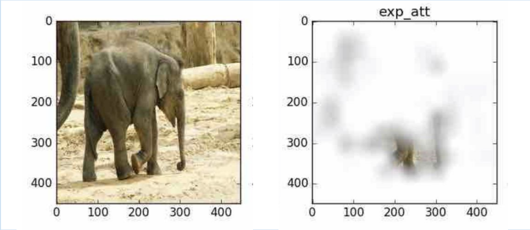
Fair

- Protected classes
- Disparate impact
- Equal opportunity



Accountable

- Explicit & implicit biases in data
- Interpretability of models & decisions



Evolving

- Social, economic, legal implications
- Dialog with society at large

Q&A

square.com