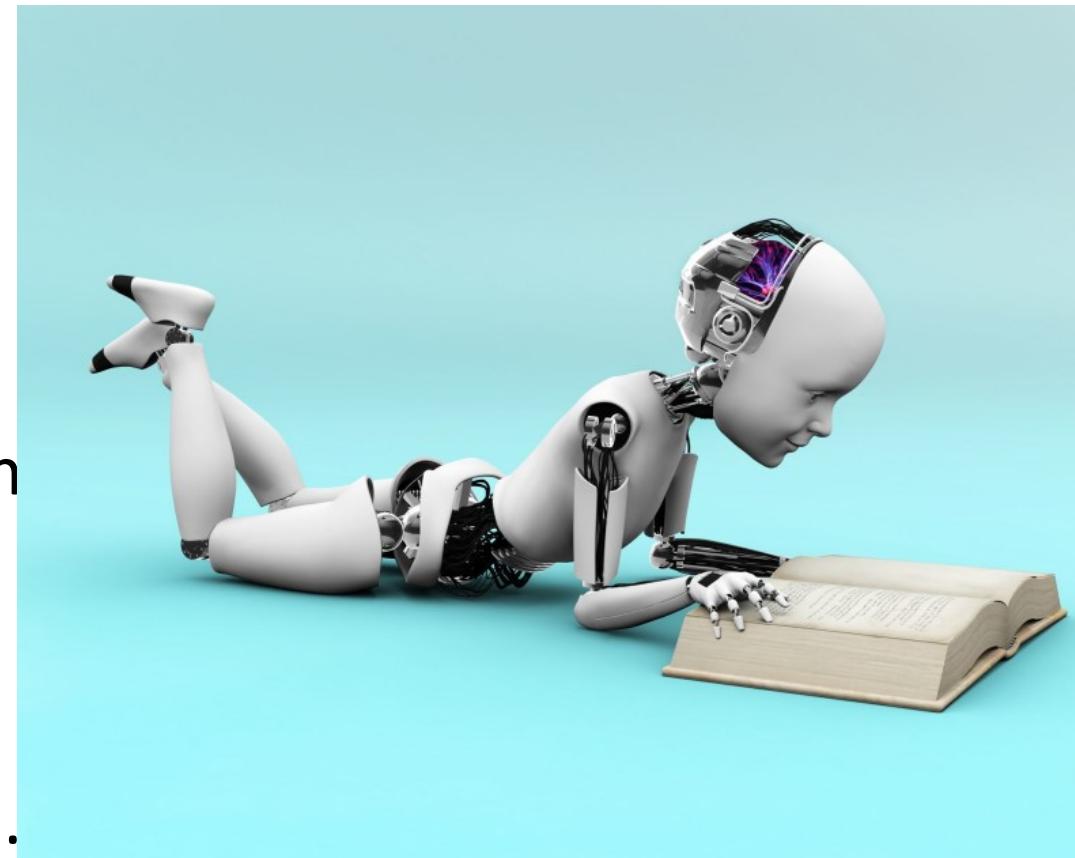


ML - Introduction

Machine Learning

- You have probably written programs (code) in the past.
- You had to program every possible situation
Think of "if – then – else" statements.
- What if machines could learn by themselves from their environment or data.



Machine Learning

- Machine learning is a method of data analysis that **automates** analytical model building.
- Using algorithms that **iteratively learn from data**, machine learning allows computers to **find hidden insights without being explicitly programmed where to look**.
- Sounds too good to be true? Well, it's already being widely used. Let's look at some examples:

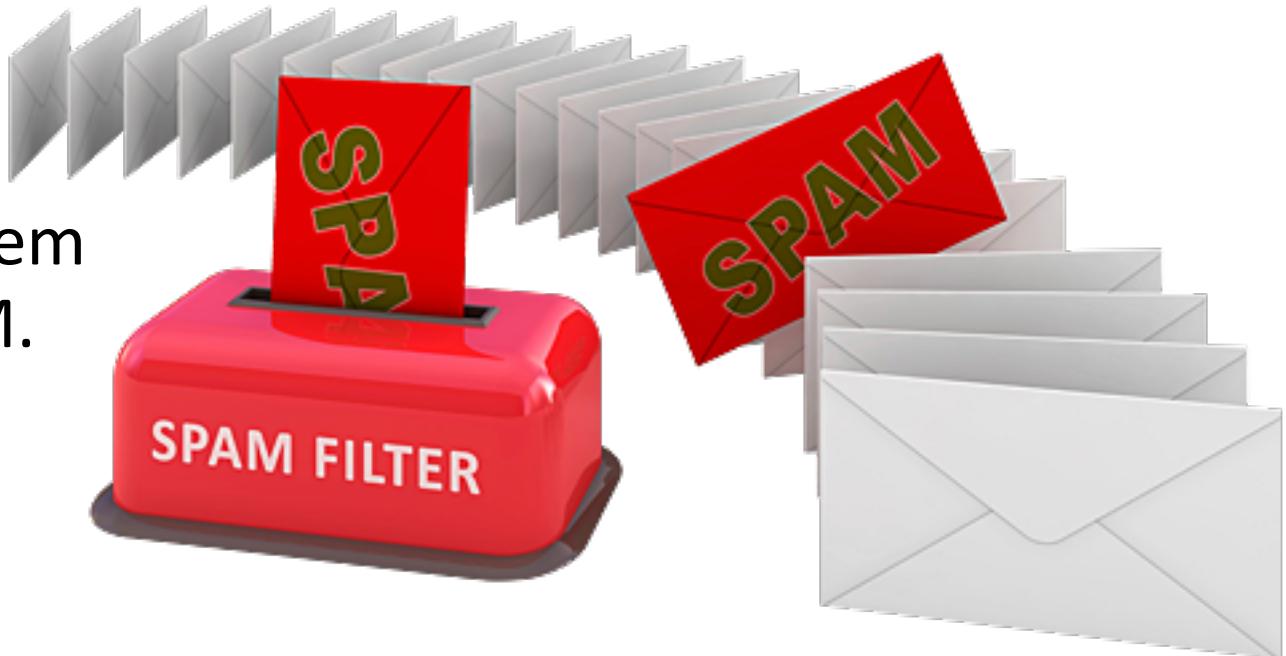
Machine Learning Examples

- Online Recommendation Systems
- System learns from your
 - purchase data
 - browsing history
 - viewing history
 - email data



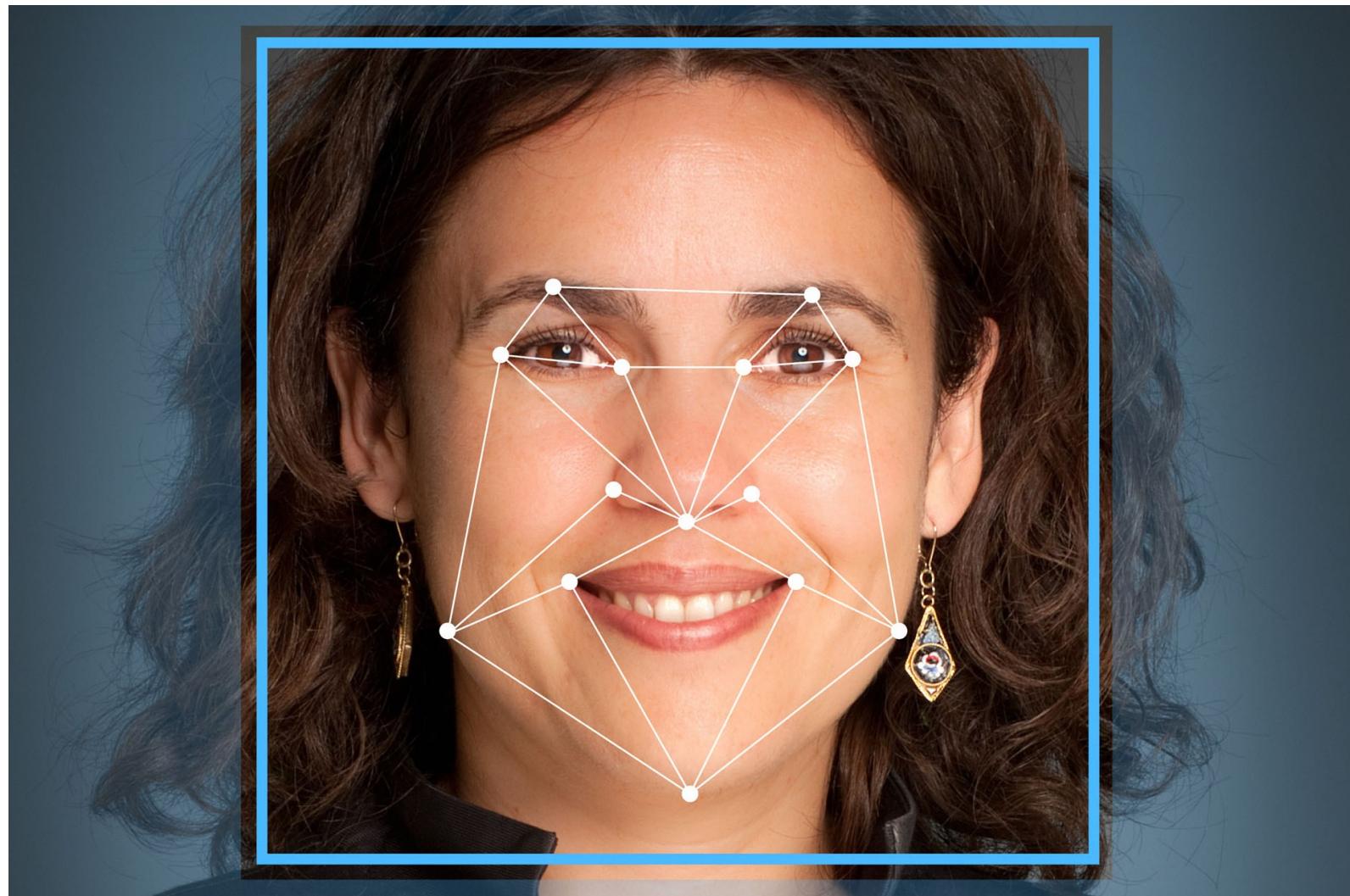
Machine Learning Examples

- Spam Filtering
- Learning from features of spam emails to classify them as either SPAM or NOT SPAM.



Machine Learning Examples

- Facial Image Detection



Famous Machine Learning Examples

- Google's Deep Mind
<https://www.youtube.com/watch?v=TnUYcTuZJpM>
- Google's Self Driving Car Project
<https://www.youtube.com/watch?v=TsaES--OTzM&t=5s>
- Six Novel ML Applications
<https://www.forbes.com/sites/85broads/2014/01/06/six-novel-machine-learning-applications/#2cf1c42a1060>

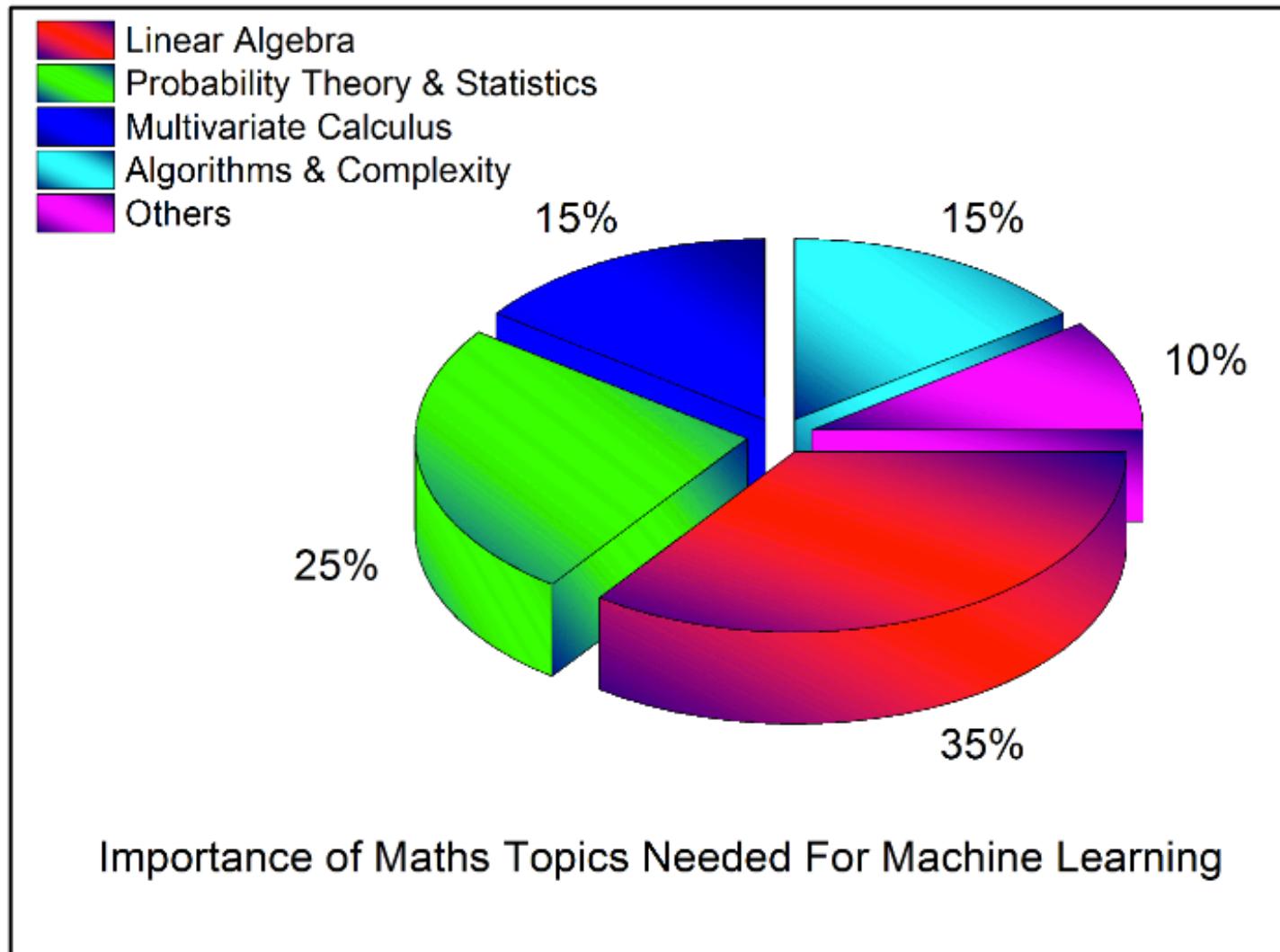
Skills Needed

I want to be a machine learning scientist -
What do I need to do / learn?

- Strong math background
- Love data
- Programming skills
- Data analysis / pattern identification
- Ability to convey results to clients / users



Math background



Skills Summary

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21th century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ★ Machine learning
- ★ Statistical modeling
- ★ Experiment design
- ★ Bayesian inference
- ★ Supervised learning: decision trees, random forests, logistic regression
- ★ Unsupervised learning: clustering, dimensionality reduction
- ★ Optimization: gradient descent and variants

DOMAIN KNOWLEDGE & SOFT SKILLS

- ★ Passionate about the business
- ★ Curious about data
- ★ Influence without authority
- ★ Hacker mindset
- ★ Problem solver
- ★ Strategic, proactive, creative, innovative and collaborative

PROGRAMMING & DATABASE

- ★ Computer science fundamentals
- ★ Scripting language e.g. Python
- ★ Statistical computing packages, e.g., R
- ★ Databases: SQL and NoSQL
- ★ Relational algebra
- ★ Parallel databases and parallel query processing
- ★ MapReduce concepts
- ★ Hadoop and Hive/Pig
- ★ Custom reducers
- ★ Experience with xaaS like AWS



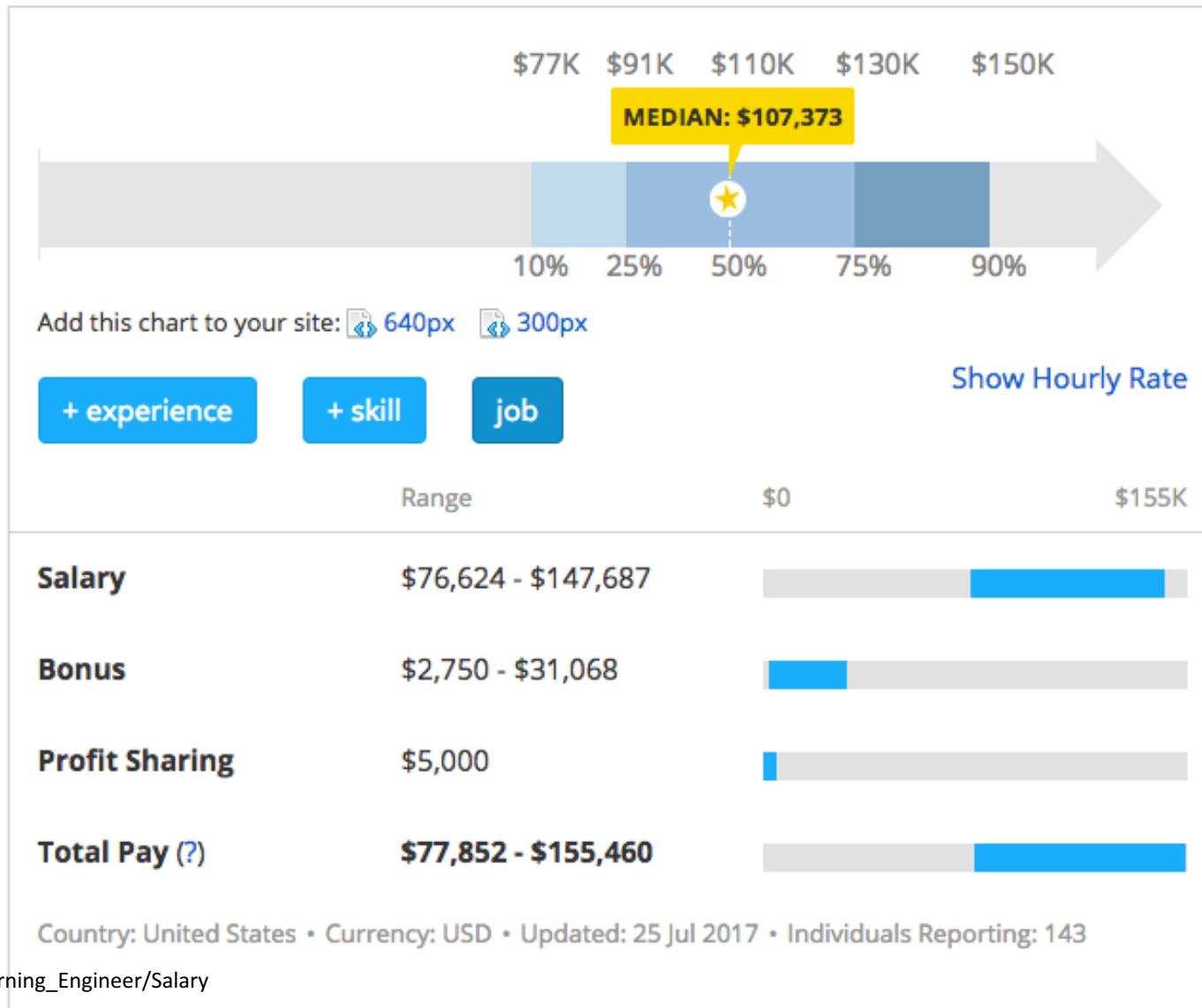
COMMUNICATION & VISUALIZATION

- ★ Able to engage with senior management
- ★ Story telling skills
- ★ Translate data-driven insights into decisions and actions
- ★ Visual art design
- ★ R packages like ggplot or lattice
- ★ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

Salary

Machine Learning Engineer Salary

A Machine Learning Engineer earns an average salary of \$100,956 per year. People in this job generally don't have more than 10 years' experience. The skills that increase pay for this job the most are Big Data Analytics and Machine Learning.



Salary

Machine Learning Engineer Salaries

43 Salaries

Updated Jul 1, 2017

National Avg

\$114,826

Min

\$85k

Max

\$137k

Dallas-Fort Worth, TX Area Avg

\$103,150*

Min

\$78k

Max

\$132k

Opportunities

- All major companies are seeking **skilled** machine learning / data scientists.
- Too many to name ☺

Let's get started

- What will this course teach me?
 - Basics of ML
 - Review of probability and statistics
 - Foundation of learning
 - Various learning instances (techniques)
 - How to apply ML to real world datasets
 - How to analyze data and results
 - How to communicate those results

