unify

A/B Testing, Statistical Experimentation and Reinforced Learning Are the Future of Product Development

Vamsi Sistla

April 1st 2021

Vamsi Sistla

About Me

- Currently –
  - Head of Data Science and Machine Learning @ Unify
- Past –
  - 15+ years of IT experience
  - 7+ Years of Data Science/Machine Learning leadership experience
  - Founded & operated 2 venture backed startups as CTO
  - Advisor/Mentor to Women In Voice (WiV) and Watson Institute

Agenda

- Module #1 - Basics of AB Testing
- Module #2 - AB Testing Quick Test Run
- Module #3 - Multi Arm Bandit & Reinforced Learning
- Module #4 - Summary and Closing Comments

Module #1
Intro AB Testing

What is AB Testing?

What is AB Testing?

Why AB Test?

What is AB Testing?

- What you know vs what you want to know
- Compare the two versions of a variable to find out which performs better in a controlled environment

Popular AB Testing Tools for Online

- Optimizely
- VWO
- Convert Experiences
- SiteSpect
- AB Tasty
- Evolv
- Google Experiments
- Qubit
- Adobe Target
- Marketing Tools With Built-In Testing

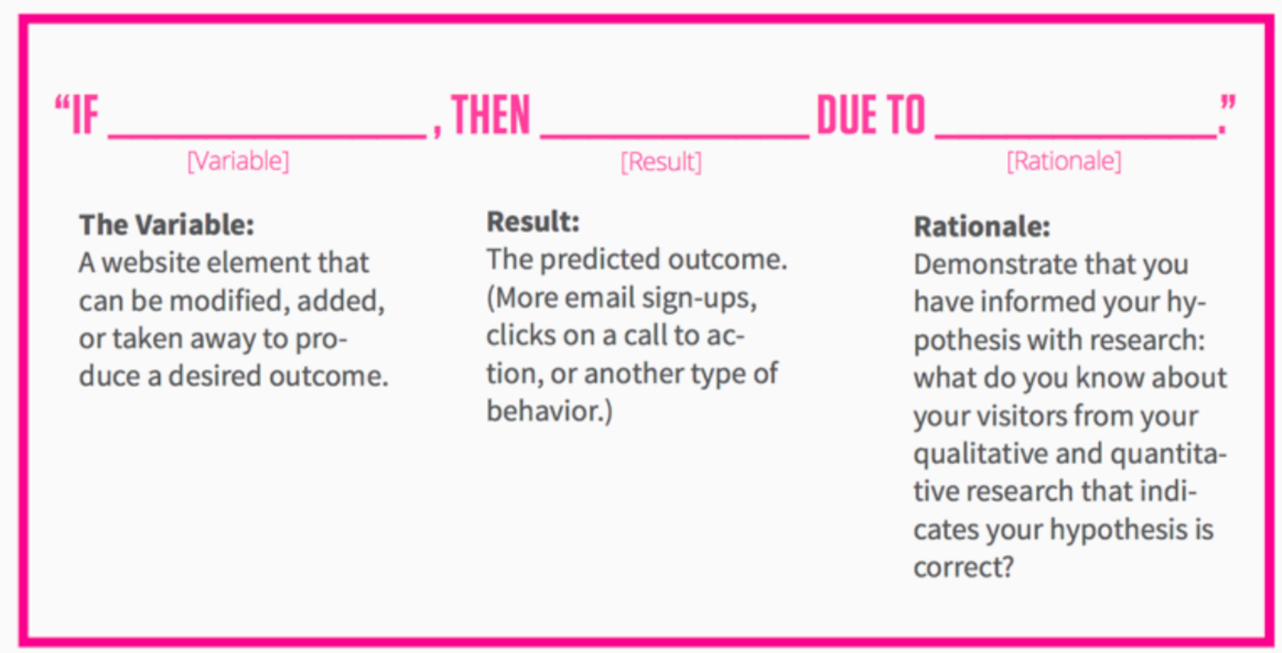
Source: conversionsciences.com

AB Testing Applications

- Web and mobile product development
- Life sciences – drug development
- Behavioral sciences and Economics
- Anywhere Causality should be assessed

Some definitions & terms

Hypothesis



Source: optimizely.com

- A **hypothesis** is a prediction that states what is being changed, what you believe the outcome will be, and why. Running the experiment will either prove or disprove your **hypothesis**
- Null Hypothesis
  - The null hypothesis is a baseline assumption that no real effect behind the data your test has produced
- Alternate Hypothesis
  - This challenges the status quo (the null hypothesis) and is basically a hypothesis that the researcher believes to be true.

Some definitions & terms

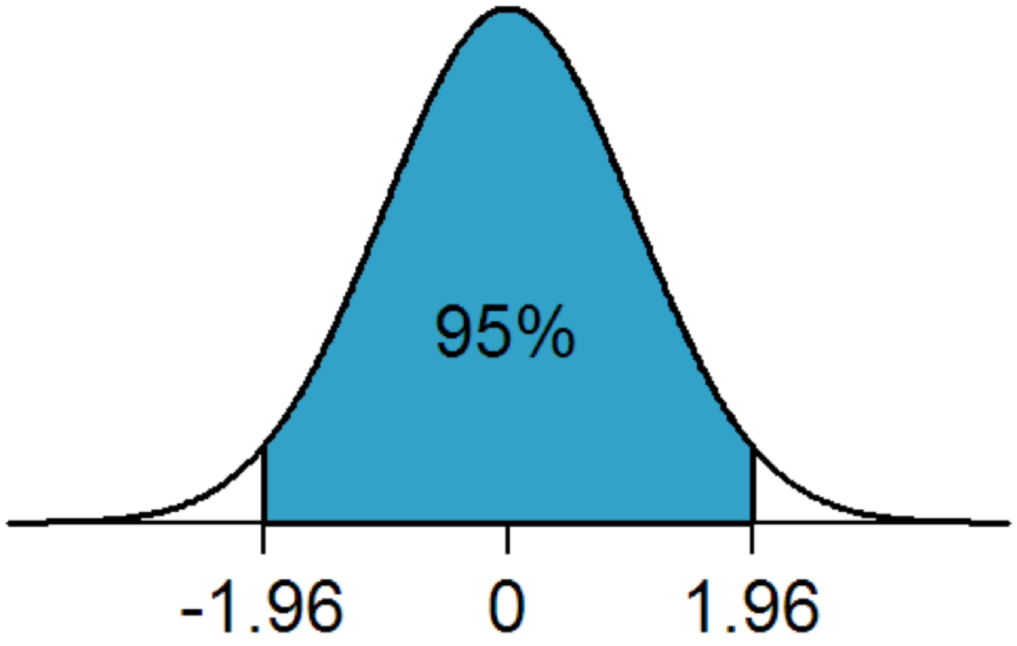
Control and Treatment

Source: Towards data science

- The **control** is simply the "Version A" of your test -- it's what you normally use as your landing page, email, call-to-action, headline, etc.
- The **treatment** is the "Version B" of your test -- it's the version that has the changes you're trying to test.
- A/B **Testing** is a way of conducting an experiment where you compare a **control** group to the performance of one or more **test** groups by randomly assigning each group a specific single-variable **treatment**.

Some definitions & terms

Statistical Significance

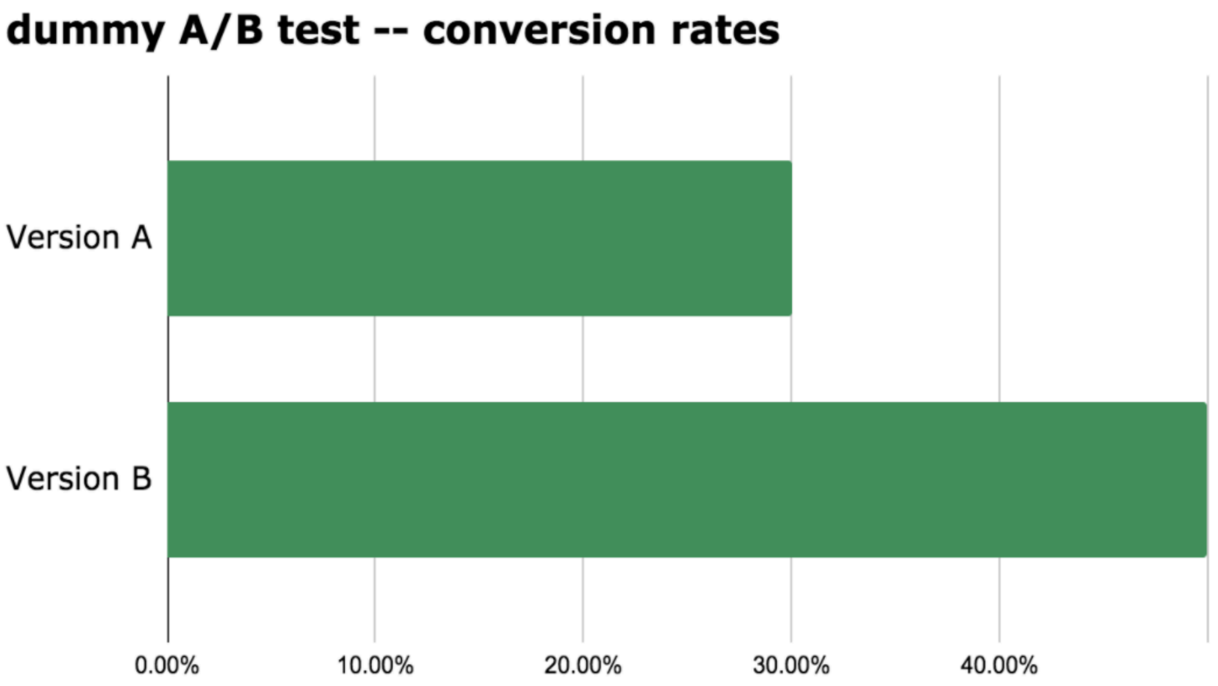


Source: Wikipedia

- Statistical significance is useful in quantifying uncertainty
- In AB testing experiments, **statistical significance** is how likely it is that the difference between your experiment's **control version** and **test version** isn't due to error or random chance.
- For example, if you run a test with a 95% significance level, you can be 95% confident that the differences are real.

Some definitions & terms

Statistical Significance



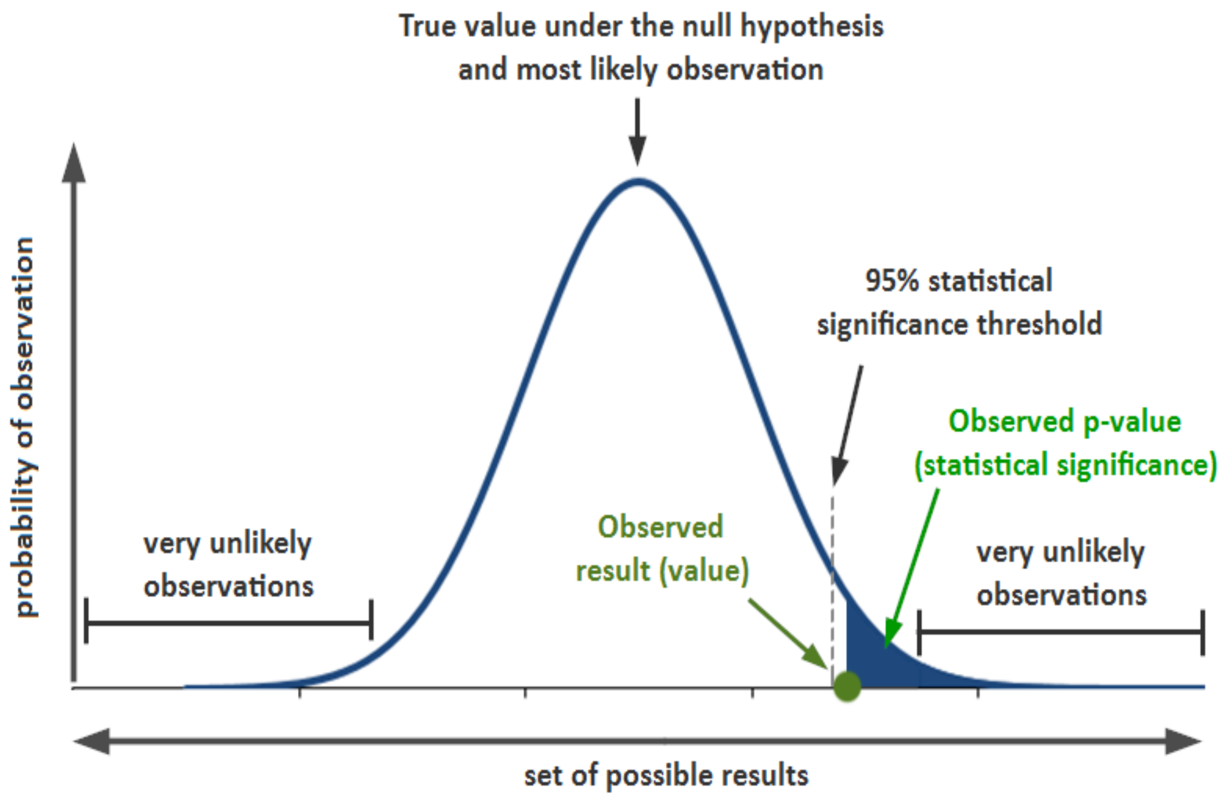
Version A: 10 users – 3 conversions – 30% conversion rate

Version B: 10 users – 5 conversions – 50% conversion rate

Source: Data36.com

Some definitions & terms

P-Value

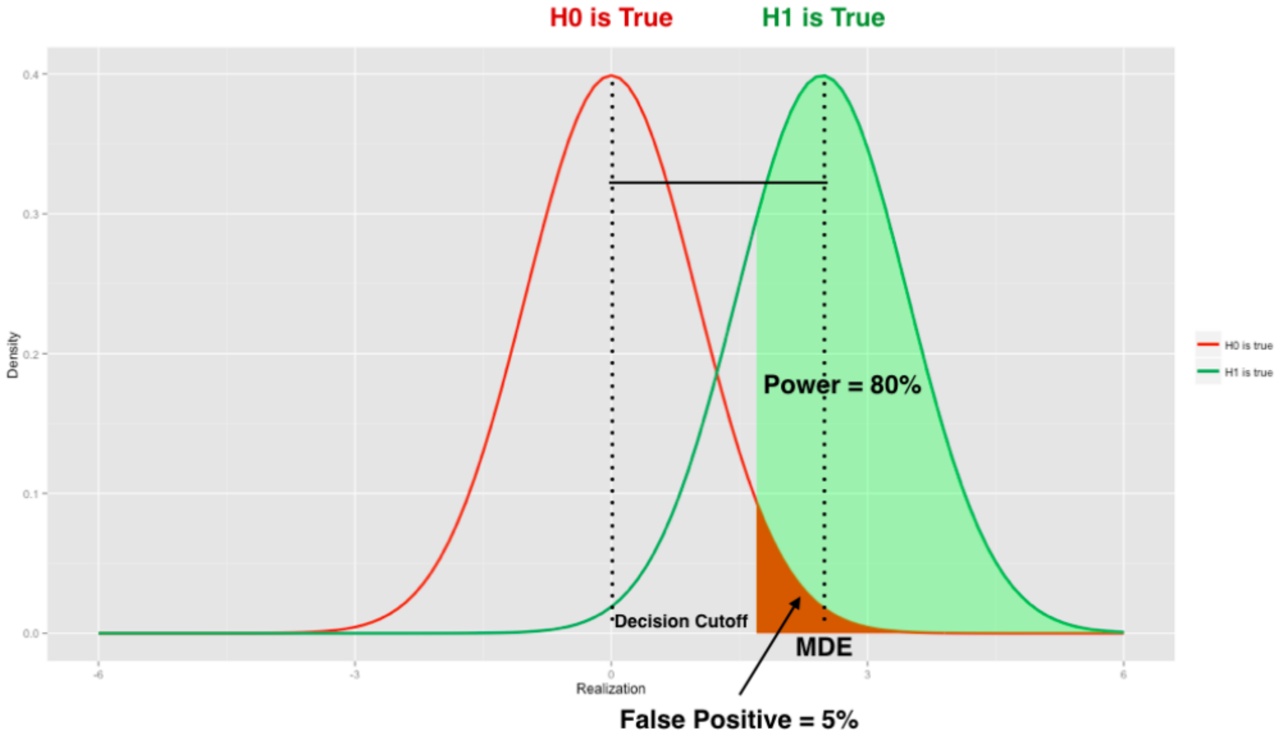


Source: SimplyPsychology

- **P-value** – a measure of evidence against the null hypothesis.
- **P-value** is created to show you the exact probability that the outcome of your **A/B test** is a result of chance.
- The level of **statistical significance** is often expressed as a **p-value between** 0 and 1. The smaller the **p-value**, the stronger the evidence that you should reject the null hypothesis. A **p-value** less than 0.05 (typically ≤ 0.05) is **statistically significant**.

Some definitions & terms

Minimal Detectable Effect

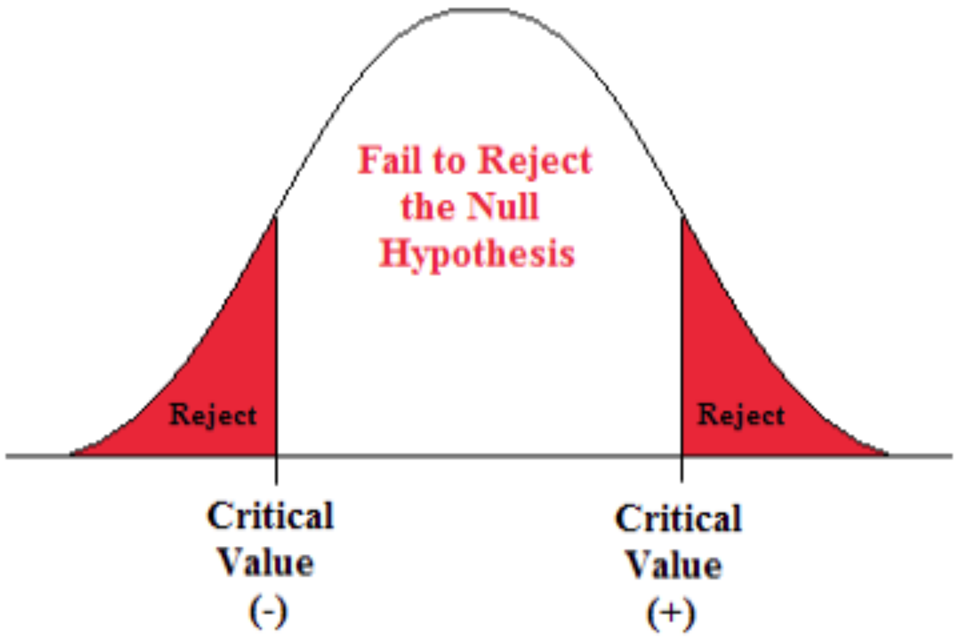


Source: Twitter blog

- **Minimal Detectable Effect (MDE)** refers to the minimal amount of change outside of error that reflects true change
- The MDE is defined as a valid change in data/outcome/result that is not due to chance.

Some definitions & terms

Rejecting the Null Hypothesis



Source: lumenlearning

- Rejecting Null Hypothesis means
  - No difference between A & B variants
  - The difference is too small to be relevant
  - There is enough difference but the **sample size** is too small to detect it.
  -

Some definitions & terms

Sampling Size

Source: Convertize

- **Sample size** calculation is based on your conversion rate, MDE, Statistical Significance and Statistical Power.
- *A sample too small* will not permit you to detect a statistically significant effect, while a *sample too large* may be a waste of limited resources.
- Sample size depends on many factors - such as what’s your definition of a winner? Which depends on your baseline When choosing sample size, consider the marginal value of added observations or, in other words, the trade-offs between accuracy and cost.

Some definitions & terms

Relationship With Sampling Size

	Relationship with Sample Size
Effect Size / Minimum Detectable Effect Size	**Inverse:** smaller the effect size, the larger the required sample size
Significance Threshold	**Inverse:** the lower the threshold, the larger the required sample size
Power / Sensitivity	**Direct:** larger the Power, the higher the required sample size

Source: Paulynn Yu blog

Some definitions & terms

Type 1 Error

Reality

	H_0 is True	H_1 is True
Do Not Reject H_0	Correct Conclusion	Type II Error
Reject H_0	Type I Error	Correct Conclusion

Source: Lumenlearning

- A type 1 error, or **alpha**, is made when H_0 is rejected, when in fact H_0 is true.
- Alpha is the probability of saying on the outcome of a test there is an effect for the manipulation, while on population level there actually is none. This is called **reliability.**

Some definitions & terms

Type 2 Error

	Reality	
	H_0 is True	H_1 is True
Do Not Reject H_0	Correct Conclusion	Type II Error
Reject H_0	Type I Error	Correct Conclusion

Source: LumenLearning

- A type 2 error, or **beta**, is made when H_0 is *not* rejected, when in fact H_0 is *not* true.
- Beta is the probability of saying on the outcome of a test there is no effect for the manipulation, while on population level there actually is. This is called **power**.
- Power is a function of alpha, sample size and effect.
- The smaller alpha, sample size or effect the smaller power is.

Some definitions & terms

Type 1 Error , Type 2 Error & Sample Size

As (Type 1 error) α increases, (Type 2 error) β decreases

As α decreases, β increases

As sample size increases (n), both **α and β** decrease

Module #2
Test Drive AB Testing

AB Testing Test Drive

Post Test Drive Comments

- With the power of Statistics and Data Sciences in general, product managers can use more tools to build superior products than ever before
- Historically, logs, data and analysis comes at the tail end of product development.

Module #3
Multi Arm Bandits & Reinforced Learning

Drawbacks with Traditional AB Testing

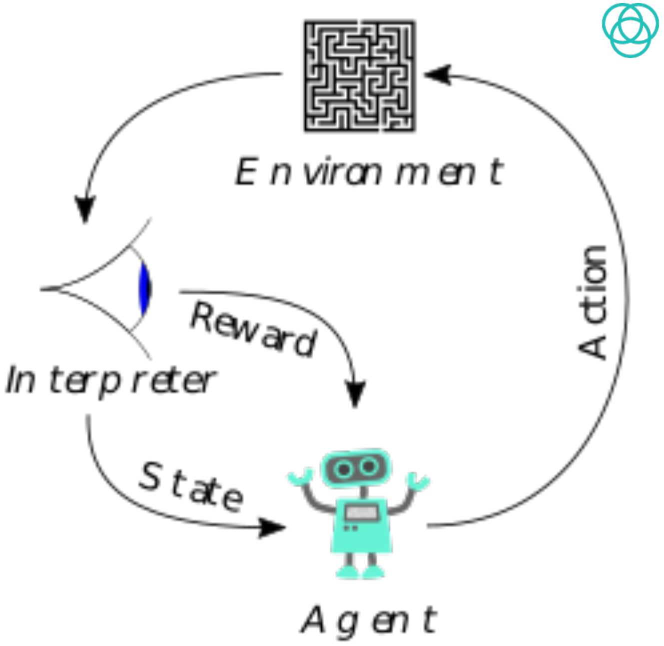
In case of traditional AB testing

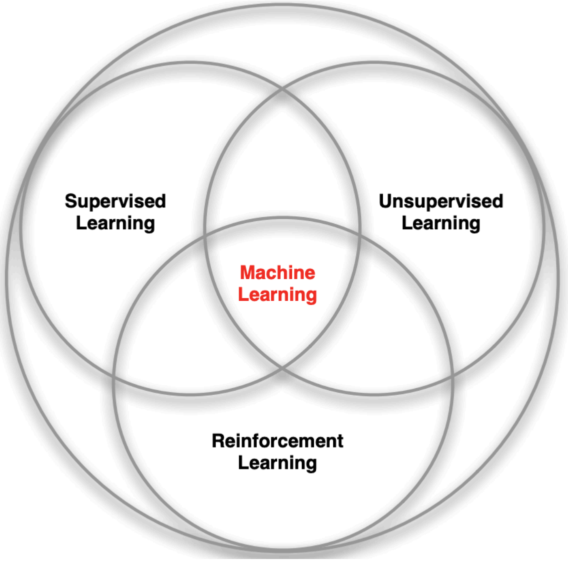
- We had to select random samples. You don’t know what should be size of your sample.
- Most of AB testing is online – it is offline – meaning you learn about which is better after the fact.
- It is time consuming to AB testing – and on top you don’t know when to stop.
- While you are at it, you might also alienate some segment of your customer who got suboptimal offering.
- Often times, traditional AB testing is limited to simple decisions – not complex systems – especially when they are sequential decisions with long term goals.

Reinforcement Learning

Overview

- Reinforcement learning (RL) is an area of ML with intelligent agents that take actions in an environment in order to maximize cumulative reward
- Reinforcement learning is one of three basic ML methods, alongside supervised and unsupervised learning.
- The focus is on finding a balance between **exploration** (of uncharted territory) and **exploitation** (of current knowledge)



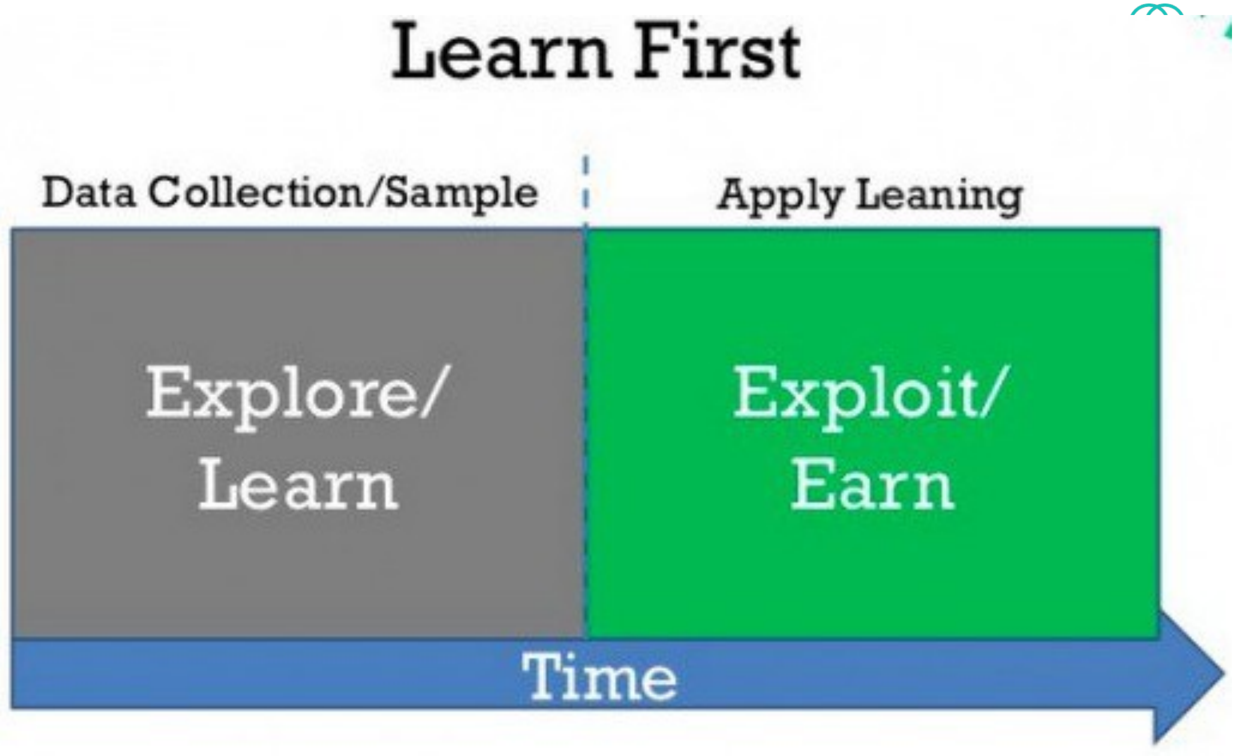


Source: Wikipedia

Reinforcement Learning

Exploration and Exploitation

- Reinforcement learning is like trial-and-error learning
- The agent should discover a good policy from its experiences of the environment without losing too much reward along the way
    - **Exploration** finds more information about the environment
    - **Exploitation** exploits known information to maximize reward
- It is usually important to explore as well as exploit



Source: Prakhar Mishra blog

Reinforcement Learning

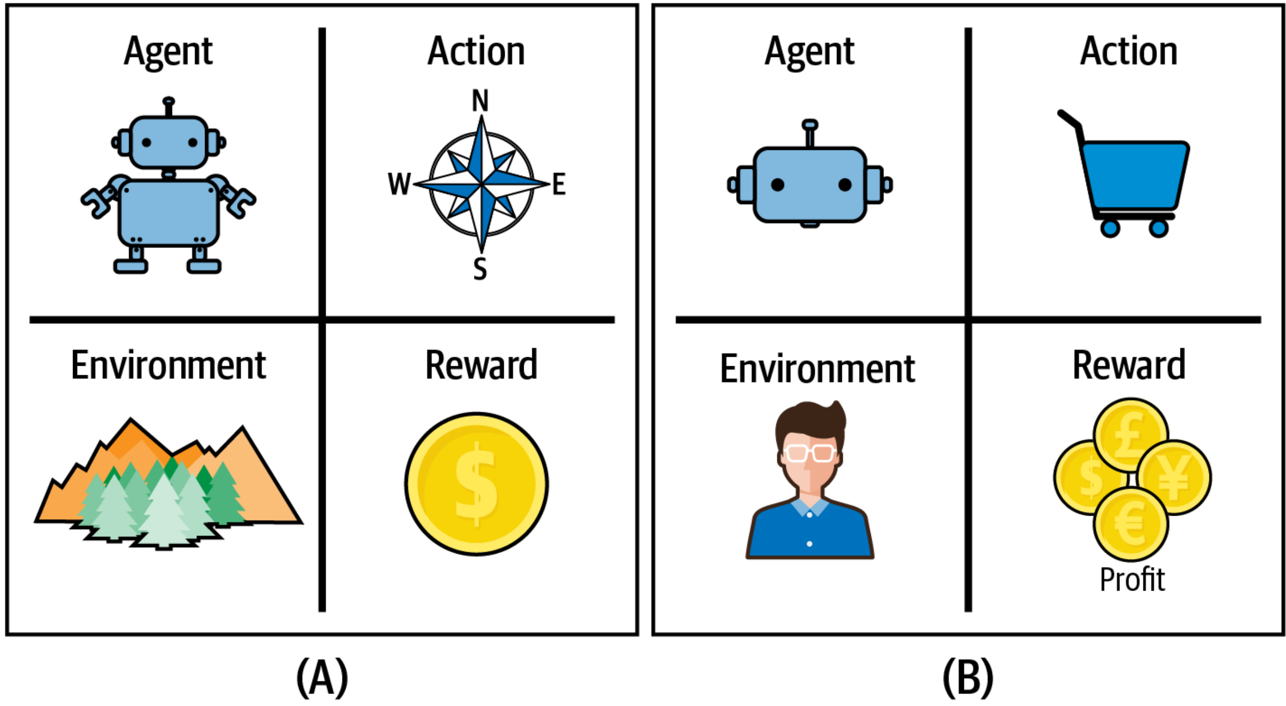
Exploration and Exploitation Examples

- Restaurant Selection
  - Exploitation Go to your favorite restaurant
  - Exploration Try a new restaurant Online Banner
- Advertisements
  - Exploitation Show the most successful advert
  - Exploration Show a different advert
- Game Playing
  - Exploitation Play the move you believe is best
  - Exploration Play an experimental move

Reinforcement Learning

Three Core Aspects

- Markov Decision Process (MDP)
  - An ***agent*** influences the observed behavior of a stochastic system (the ***environment***) by choosing *actions*. The goal is to choose the set of **actions** that allow the system to behave in an optimal way, as defined by some success criterion (the ***reward***).
- Dynamic Programming
  - is a collection of algorithms that can solve a problem with *perfect model of the environment and where an agent can only take discrete actions.*
- Monte Carlo
  - Monte Carlo Simulations Are Like Unit Tests for Bandit Algorithms



Example (A) shows a robot that intends to move through a maze to collect a coin.
Example (B) shows an ecommerce application that automatically adds products to users' baskets, to maximize profit.

Source: Reinforcement Learning by Phil Winder

RL – Some Examples

- Cathy WU – UC Berkeley/MIT – Built a traffic management scenario (Deep RL) -
  - https://sites.google.com/view/ieee-tro-flow/home

RL – Multi Arm Bandit Algorithms

What is MAB?

Multi-armed bandits (MAB) is a type of RL.

Multi-armed bandits extend RL by ignoring the state and try to balance between exploration and exploitation.

Applications

- Website design
- Clinical trials
- Marketing & Recommender Systems

A multi-armed bandit solution is a 'smarter' or more complex version of A/B testing that uses machine learning algorithms

- Optimizely

Bandits vs ML

How are Bandits different from typical ML algorithms?

Unlike standard machine learning tools, bandit algorithms aren’t simply black-box functions you can call to process the data you.

Bandit algorithms have to actively select which data you should acquire and analyze that data in real-time.

Indeed, bandit algorithms support two types of learning that are not present in standard ML examples:

- ***active learning***, which refers to algorithms that actively select which data they should receive; and
- ***online learning***, which refers to algorithms that analyze data in real-time and provide results on the fly.

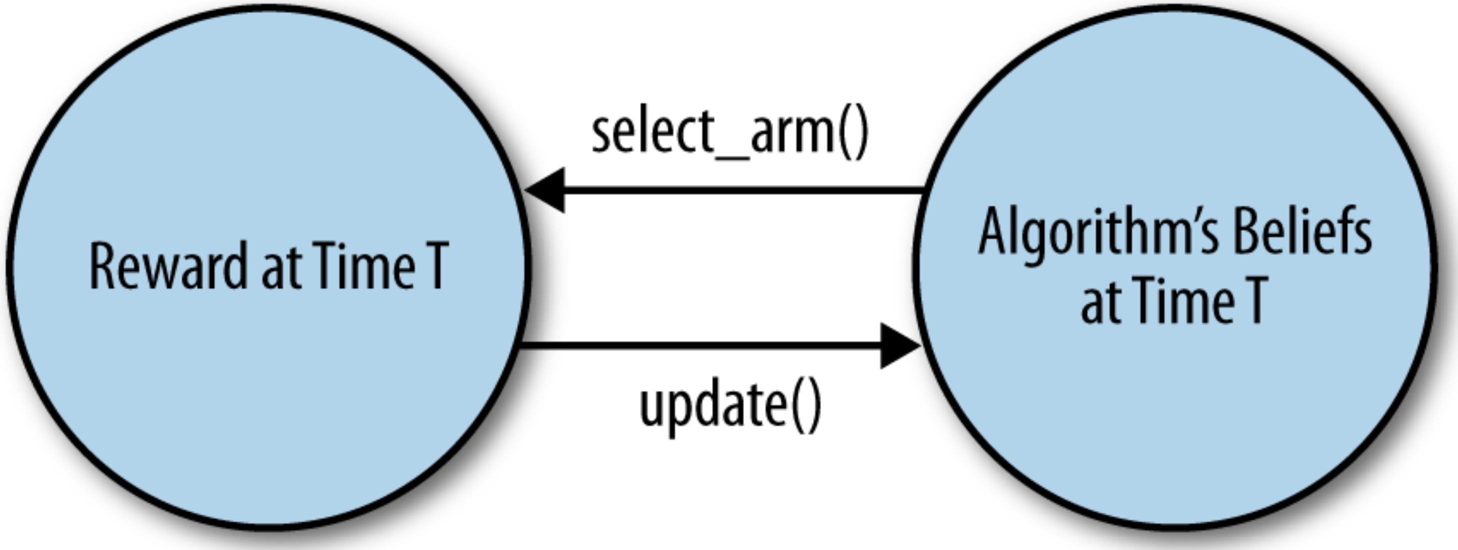
RL – Multi Arm Bandit Algorithms

So, how to use MAB to do AB Testing?

- Reward Engineering
  - To quantify performance, the result of an action must be measurable. In RL, this is the purpose of the reward.
  - There are many ways to quantify the performance.
- The Actions
  - Bandit approach to optimize your returns.
  - Compromise between exploitation and exploration
- The environment
  - If you have prior data - A simulated environment that can mimic a real world scenario?
  - If you don’t have prior data – then explore the possibility of using a segment of the real world environment

RL – Multi Arm Bandit Algorithms

Rewards



- Behavior of Algorithm depends on the data
- What data the algorithm sees depends on its behavior

RL – Multi Arm Bandit Algorithms

Rewards

- Average of all the rewards
    -
- Iterative Online Approach
- The environment

$$r^{avg}(a) \doteq \frac{1}{N(a)} \sum_{i=1}^{N(a)} r(a)_i = \frac{r_1 + r_2 + \cdots + r_{N(a)}}{N(a)}$$

RL – Multi Arm Bandit Algorithms

Actions

input: a exploration probability $0 \leq \epsilon \leq 1$,

Initialize $r^{avg}(a) \leftarrow 0, N(a) \leftarrow 0$, for each $a \in \mathcal{A}$

loop for ever:

$$a \leftarrow \begin{cases} \underset{a_s \in \mathcal{A}(s)}{\text{argmax}}\ r^{avg}(a_s) & \text{with probability } 1 - \epsilon \text{, breaking ties randomly} \\ \text{random } a & \text{with probability } \epsilon \end{cases}$$

Present action a in the environment and receive reward r

$N(a) \leftarrow N(a) + 1$

$$r^{avg}(a) \leftarrow r^{avg}(a) + \frac{1}{N(A)}\left[r - r^{avg}(a)\right]$$

- Average of all the rewards
    -
- Iterative Online Approach
- The environment

RL – Multi Arm Bandit Algorithms

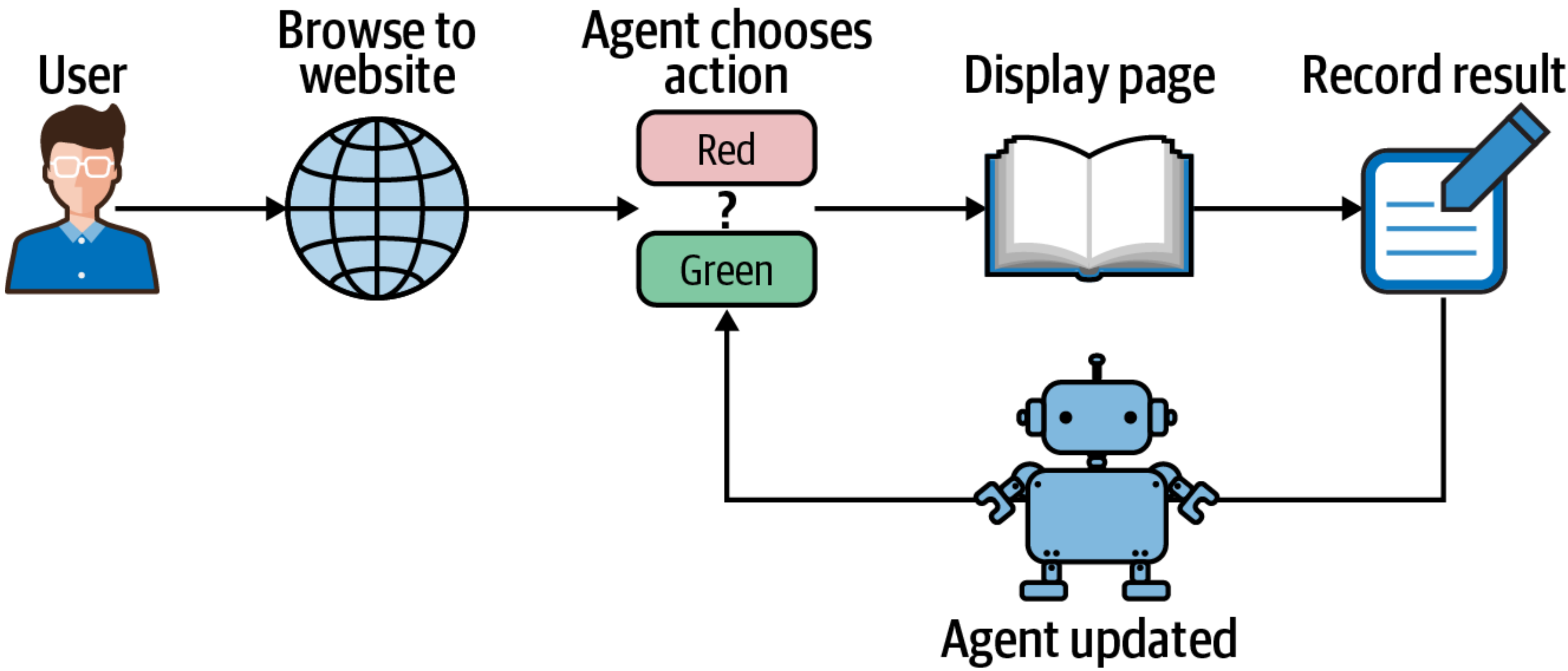
Environment & running the experiment

- Average of all the rewards
    -
- Iterative Online Approach
- The environment

function $\textsc{Environment}(a, p(a))$

input: The action, a, and its probability, $p(a)$

output: $r \leftarrow 1$ with probability $p(a)$, else 0.



Bandit Algorithms

Types of Bandit Algorithms

- Stochastic Bandits
  - Ex – Epsilon Greedy, Uniform Exploration, Adaptive Exploration
- Bayesian Bandits
  - Ex – Thompson Sampling
- Lipschitz Bandits
  - Ex – Continuum Armed
- Adversarial Bandits
- Linear Bandits
- Contextual Bandits

Bandit Algorithms

Epsilon Greedy - Definition

- The epsilon-Greedy algorithm is one of the easiest bandit algorithms tries to be fair to the two opposite goals of exploration and exploitation by using a mechanism of randomly choosing one of the options.
- That randomness is set by epsilon - ϵ
- It works by oscillating between (A) exploiting the best option that it currently knows about and (B) exploring at random among all of the options available to it

Arm

- Arm is this algorithm is option. So, in case of A/B testing, it would be 2 options or arms.

Bandit Algorithms

Epsilon Greedy

With probability $1 - \epsilon$, the algorithm ***exploits*** the best known option.

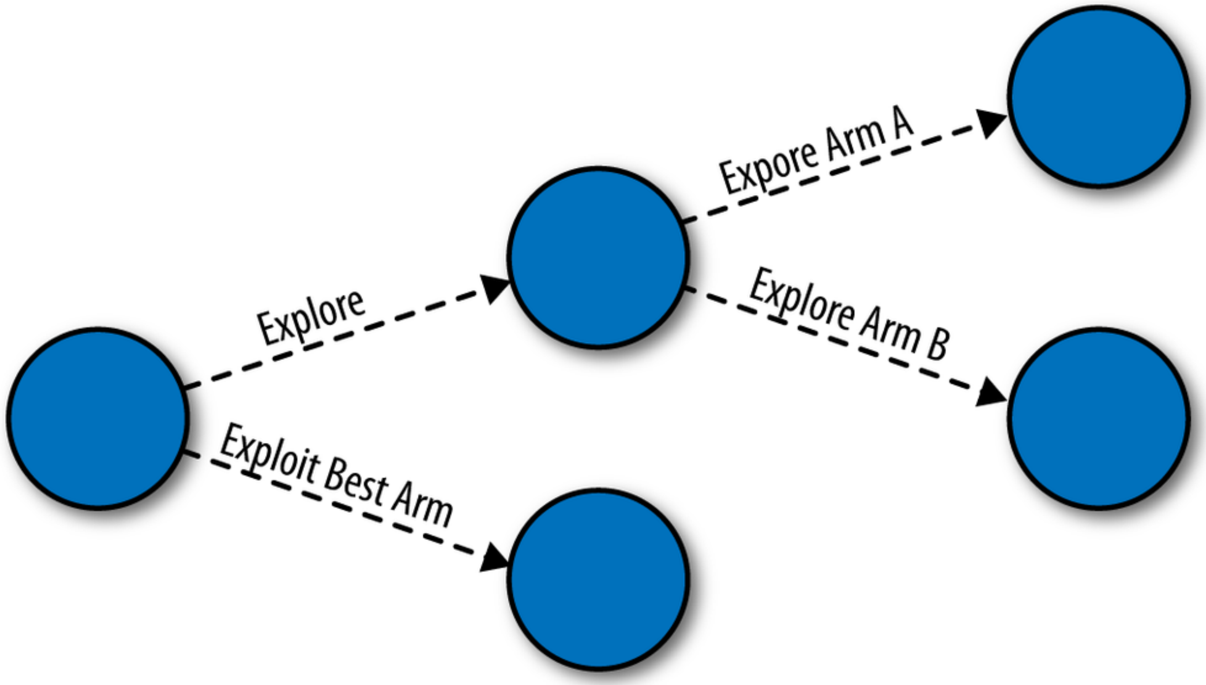
With probability $\epsilon / 2$, the algorithm ***explores*** the best known option.

With probability $\epsilon / 3$, the algorithm ***explores*** the worst known option.

Bandit Algorithm Test Drive

Epsilon Greedy Algorithm

- Greedy Algorithm that adds some randomness when deciding: Instead of picking always the best available option, randomly explore options with a probability = ϵ or pick the best option with a probability = 1 - ϵ.
- ϵ is a hyper-parameter that needs to be tuned based on the experiment, i.e. there is no value that works best on all experiments

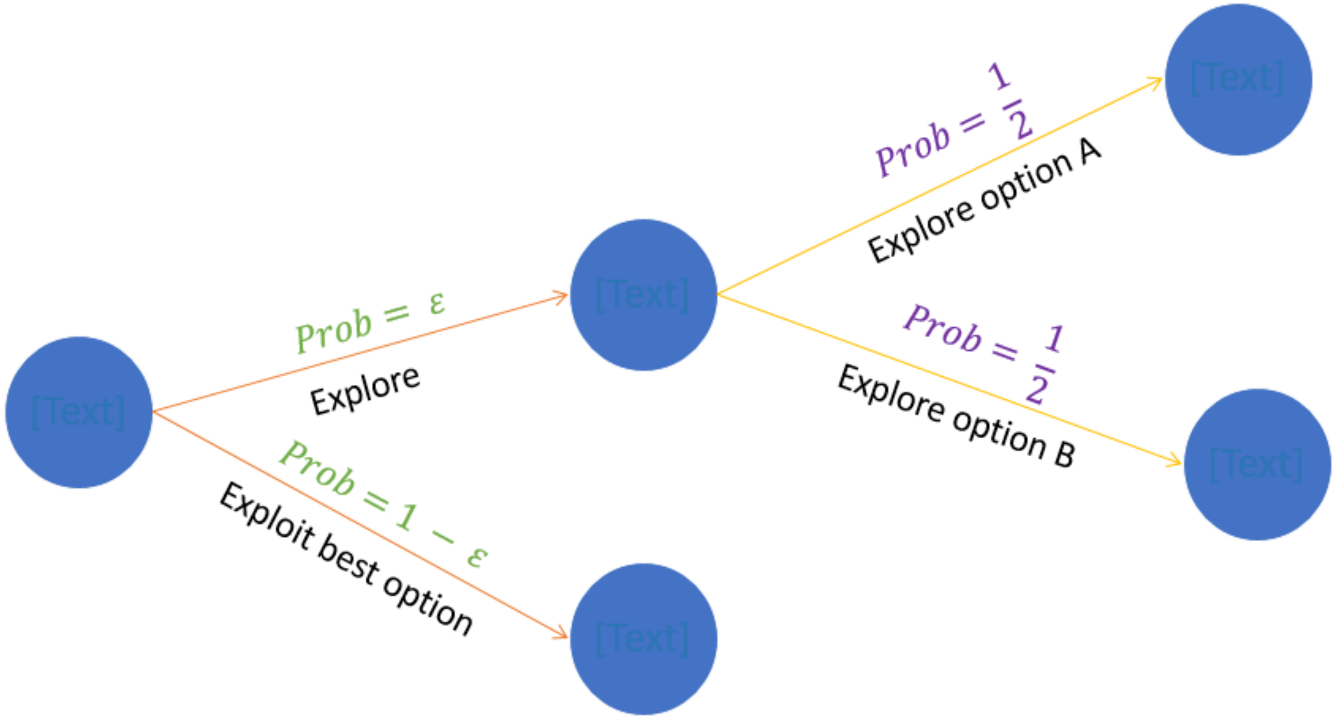


The epsilon-Greedy Algorithm

Bandit Algorithm Test Drive

Multi Arm Bandits - Demo 1

- In this demo, author used coin toss as the task.



Source: imaddabbura.github.io

Bandit Algorithm Test Drive

- Some of the code from This repository was forked from John Myles White's "BanditsBook" repository.
    - https://pypi.org/project/banditsbook/
    -

- Some from Dr. Phil Winder’s work in RL Book
    - https://rl-book.com/
- Other comments

Module #4
Wrap up and Summary

Epsilon Greedy Takeaways

- The epsilon-Greedy algorithm does eventually figure out which arm is best no matter how epsilon is set. But the length of time required to figure our which arm is best depends a lot on the value of epsilon.
- What’s appropriate for you depends on how long you intend to run your algorithm and how different the arms you’re testing will be

Key Takeaways

- AB Testing is a broad and deep topics. PDs should use this as an arsenal to make data driven decisions when building products
- Traditional AB testing worked wonders for websites and mobile apps but comes with many challenges.
  - Especially ideal non sequential and binary decisions.
  - Risk of alienating your users and prospective customers
- Product development should start using RL algorithms to learn and iterate
  - Multi Arm Bandit Algorithms are super powerful

Blogs & Articles

- Power, minimal detectable effect, and bucket size estimation in A/B tests - **https://blog.twitter.com/engineering/en_us/a/2016/power-minimal-detectable-effect-and-bucket-size-estimation-in-ab-tests.html**
- Understanding Power Analysis in AB Testing - **https://towardsdatascience.com/understanding-power-analysis-in-ab-testing-14808e8a1554**
- Imad Dabbur Blog on Epsilon Greedy Algorithm -**https://imaddabbura.github.io/**
  - He implemented based on the book - Bandit Algorithms for Website Optimization
- Google Colab - ### For those of you want a refresher for Google Colab, here's a source for you to get started - https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c

- Another blog - https://github.com/raffg/multi_armed_bandit
- Code & Files
  - I have ran the demos using Google Collab. So, you don't need to do any special setup on your laptop. If you have Gmail a/c, you are off to races.
  - Location of these files –