

# ML @ Snackable

Alexander Tkachenko

[SNACK]ABLE

# Outline

- What we do and why
- Spoken language
- ML project lifecycle

# Spoken Word Content



interviews



social video



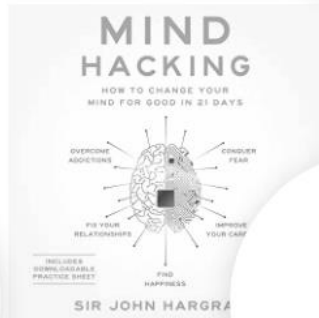
radio



youtube



podcasts



audiobooks

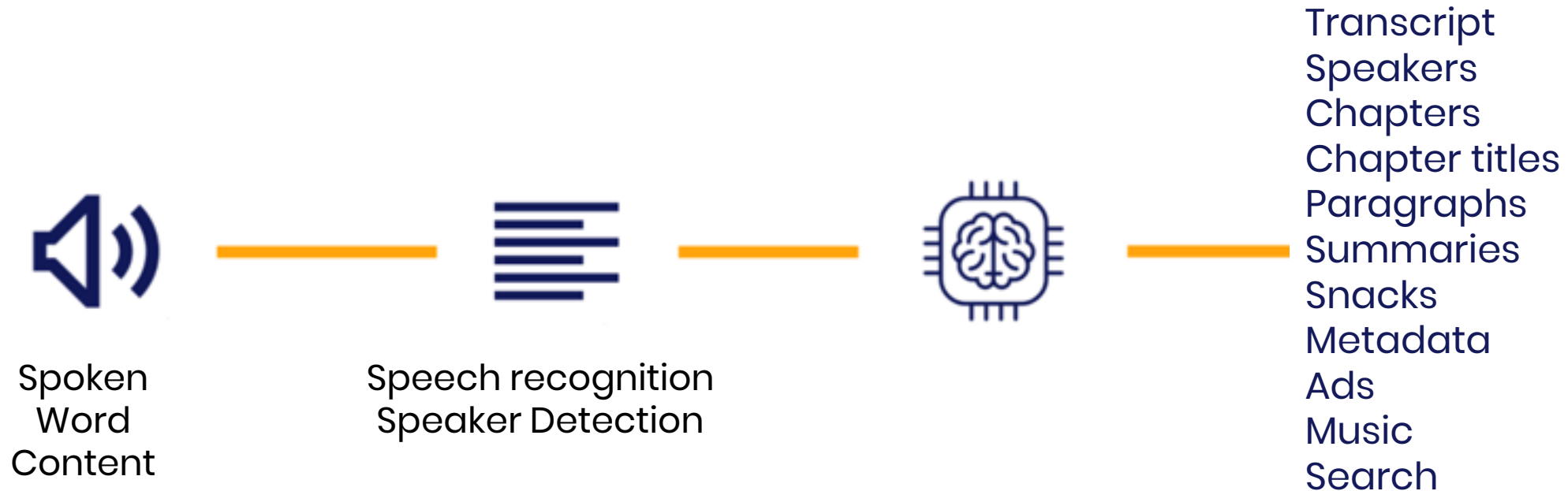


speeches



tv

# Snackable AI Engine



# Rich Transcript

The screenshot displays a video player interface with a rich transcript. On the left, a 'Chapters' sidebar lists seven segments: 1 (00:05 - 03:55) 'The aggregation of marginal gains', 2 (03:56 - 09:14) 'Building the habits that you need to get results', 3 (09:15 - 13:24) 'How can we change people's habits without asking them to do anything', 4 (13:25 - 17:11) 'The 2 minute rule for productivity', 5 (17:13 - 18:49) 'Optimizing for the starting line', 6 (18:50 - 21:05) 'How to be a better comedian', and 7 (21:07 - 24:35) 'Why habits are so important'. The main content area shows the second chapter, 'Building the habits that you need to get results', with a play button and the name 'James Clear'. Below the name are tags: 'implementation intention', 'habit', 'motivation', 'exercise', 'Strategy', and 'behavior'. The transcript text reads: 'And so throughout the rest of this presentation, I want to talk about how we can do that. Today, I'm going to teach you how to build the habits that you need to get the results that you want. And in order to do this, I'm going to take you through a framework for building better habits. And then I'm also going to share a personal example of how I use this. So my writing habit. If you don't know, I write a James Clear dot com. I write about how to build better habits, improve performance, and generally live better. Over a million people visit the site each month. There's over 400,000 subscribers on the weekly email newsletter, and it all came out of the simple writing habit.' Below this is an 'Edit text' option and another play button. To the right is a video player showing a man on stage with a screen behind him that says 'Good habits make time your ally. Bad habits make time your enemy.' and 'JamesClear.com'. The video player has a progress bar at the bottom showing 03:58 / 24:36. At the bottom of the entire interface is a timeline with markers for 04:55, 09:50, 14:46, 19:41, and 24:36, and a small '2' icon indicating the current chapter.

**Chapters**

- 00:05 - 03:55  
The aggregation of marginal gains
- 03:56 - 09:14  
Building the habits that you need to get results
- 09:15 - 13:24  
How can we change people's habits without asking them to do anything
- 13:25 - 17:11  
The 2 minute rule for productivity
- 17:13 - 18:49  
Optimizing for the starting line
- 18:50 - 21:05  
How to be a better comedian
- 21:07 - 24:35  
Why habits are so important

**2 Building the habits that you need to get results**

implementation intention habit motivation exercise Strategy behavior

James Clear + Add to Snack

And so throughout the rest of this presentation, I want to talk about how we can do that. Today, I'm going to teach you how to build the habits that you need to get the results that you want. And in order to do this, I'm going to take you through a framework for building better habits. And then I'm also going to share a personal example of how I use this. So my writing habit. If you don't know, I write a James Clear dot com. I write about how to build better habits, improve performance, and generally live better. Over a million people visit the site each month. There's over 400,000 subscribers on the weekly email newsletter, and it all came out of the simple writing habit.

Edit text

+ Add to Snack

So for the rest of this talk, there are four stages. Have information. I'm going to take you through each of those four. All right. So the four stages are noticing, wanting, doing and liking, noticing, wanting, doing and liking. You cannot perform a habit or take an action if you do not notice something. I need to see a coffee cup sitting on the side in order to pick it

Good habits make time your ally.  
Bad habits make time your enemy.

JamesClear.com

03:58 5 24:36

04:55 09:50 14:46 19:41 24:36

# Snacks

Snacks 5 + New Snack

Trimming audio i Add/remove segments

Small habits and little choices are the key to success  
May 10, 2022 - 56s

How to bring a reward into the present moment  
May 10, 2022 - 57s

A study on exercise and motivation  
May 10, 2022 - 58s

last five. Now I've got two British cyclists, but it was at the Olympics in London in 2012 when this kind of strategy really came to fruition. They won 70% of the gold medals available. And so this idea that small improvements, tiny habits, little choices are not just a cherry on top of our performance, not just like a nice thing to have, but actually can be the key that unlocks significant success. That's an idea that I want us to carry with us as we go through the rest of this presentation.

And one way to think about it is just kind of basic math. Like if you just look at the numbers, if you were able to improve by 1% each day for an entire year and those gains compound, you would end up 37 times better at the end of the year. And if you were to get 1% worse, you would whittle yourself almost all the way down to zero. And what we what's interesting here is that everybody wants a transformation, right? Everybody wants a radical improvement on rapid success. But we fail to realize that small habits and little choices are transforming us every day already, that these times when you make a choice is slightly better, slightly worse, a little mistake or a small error 1% better or 1% worse that these things compound over time. And habits are the compound interest of self-improvement. And so if you can learn to master those, then you can make time work for you rather than against you. Right? Good habits make time your ally. Bad habits make time your enemy.

00:00 5 5 00:56

Small habits and little choices are the key to success  
Duplicate Snack Delete Snack  
Duration 56s  
Modified on May 10, 2022

# Use Cases

- Promotion: sharing snacks / audiograms in social media
- Rich transcript pages
- Improved navigation: chapters on YouTube
- Search

# Spoken Language

- Less structured, less grammatical
- Disfluencies
  - Uh, um, word repetitions, etc.
- Multiple speakers
- Low information density
- Variety of topics



# What it means

- ASR is often not perfect:
  - word transcription (especially entities)
  - sentence boundaries and punctuation.
- Models trained on written text can have poor results.

# Solution

- Fix ASR using custom vocabulary
  - Analyze metadata (titles, rss, descriptions, etc)
  - Video - OCR
- Fine-tune text-based model
- Use additional signals from audio/video
  - Scene detection
  - Acoustics features

# ML Project Lifecycle



# ML Project Lifecycle



- Define a Problem
- Data
- Resources
- KPI / Metrics

# ML Course/Competition/Research

- Problem
  - is well-defined in terms of X and Y.
  - the goal is to beat a baseline.
- Data:
  - is provided by organizer
- KPIs/Metrics
  - defined by organizer
- Resources
  - personal time, computational power

# In Industry

- ~~Let's try something cool~~
- High-level goals:
  - Improve a product
  - Gain competitive advantage
  - Optimize a process
- Estimate cost/risk/reward/resources
- KPIs/Metrics
  - ML metrics vs. business metrics
- ML vs. business side.

# ML Project Lifecycle



Do we have data?

## **No:**

- Where to get data?
  - use public data
  - buy
  - annotate in-house/outsource
- How much data is needed?

## **Yes:**

- Is our data usable?
  - exploratory analysis
  - what's missing?
- Build required sample:
  - export, clean, normalize, transform.

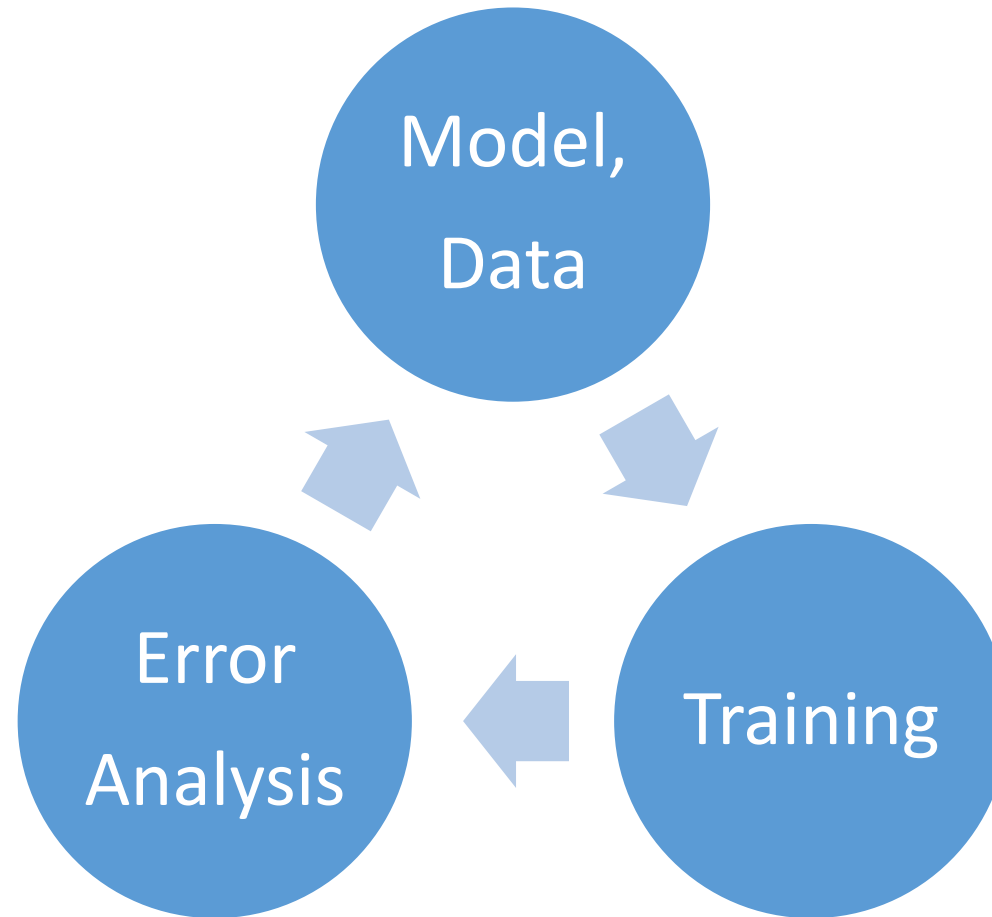
# ML Project Lifecycle



- Writing code
- Training model
- Error analysis

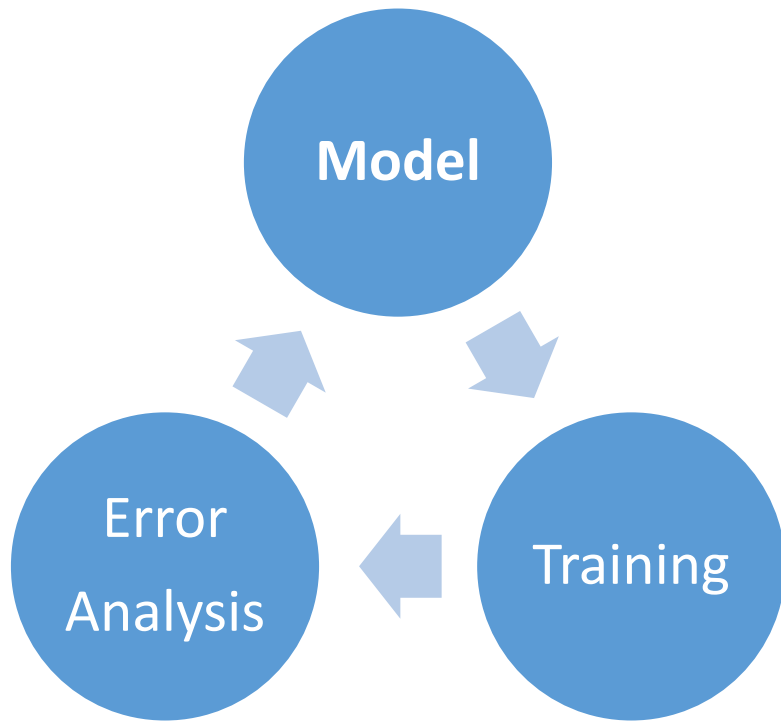


# Iterative Model Development

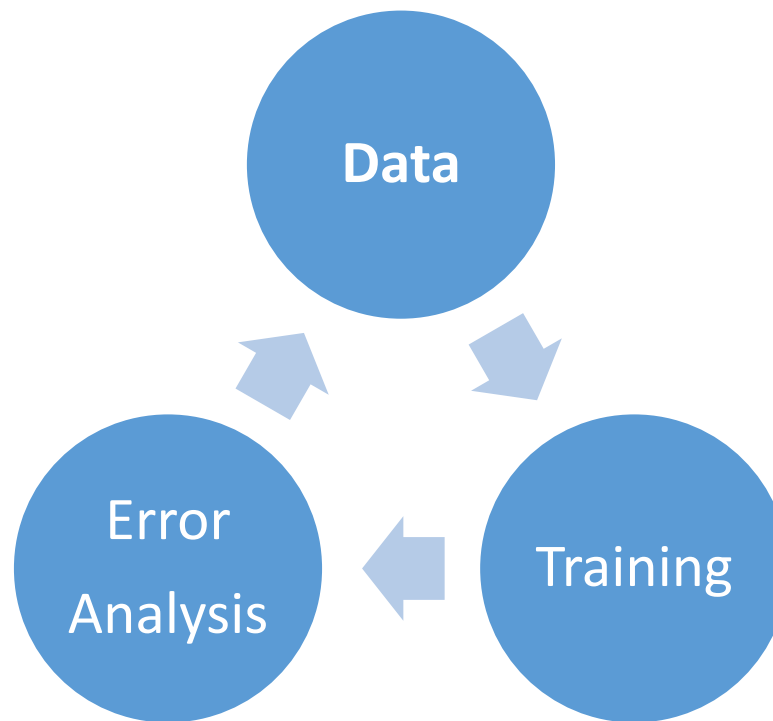


# Data-centric Approach

Research



Industry



# Reproducibility

- Documentation is crucial
- Version control
  - Data
  - Code
  - Model weights
  - Performance metrics
- Tools
  - Code: Github
  - Data/metrics: dvc, github lfs, etc...
- It's easy!

# Example: Chapter Segmentation



Phase 1 – get started quickly

- Simple unsupervised model
- “Simple” data: news, podcasts

# Model Development

- Tuning parameters:
  - Type of embeddings?
  - Text preprocessing (stopwords, lemmatization, weighting)
  - How many chapters?
  - ...

# Error Analysis

- Good:
  - detecting topic drift in information-dense content (e.g. news)
- Bad:
  - Fails with less informative content (e.g. conversations)
  - Can't learn useful signals:
    - “Let's now move to the next question...”
    - “Second, I want to talk about ....”
  - Identified topic boundaries are inexact.

# Chapter Segmentation : Phase 2



- Data: podcasts, talk shows, interviews, webinars.
- More powerful supervised model.

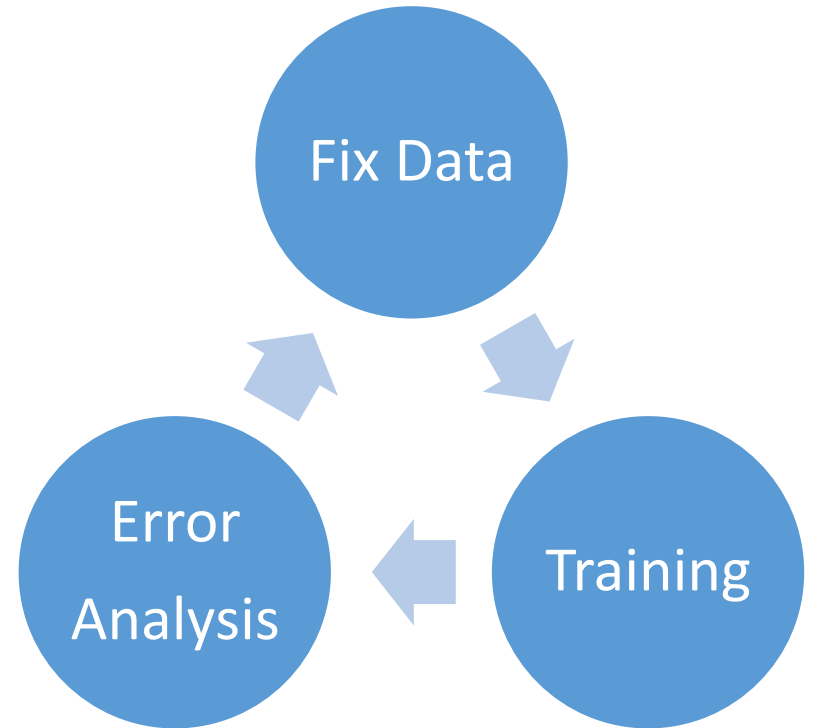
# Annotation project

- Data collection
  - Decide which types of content are important
  - Create a representative and balanced sample
- Define what is “chapter”
  - and chapter categories (ad, music, intro, outro, presentation, sub-topic, etc, ...)
- Implement a simple annotation tool.
- Hire annotators.

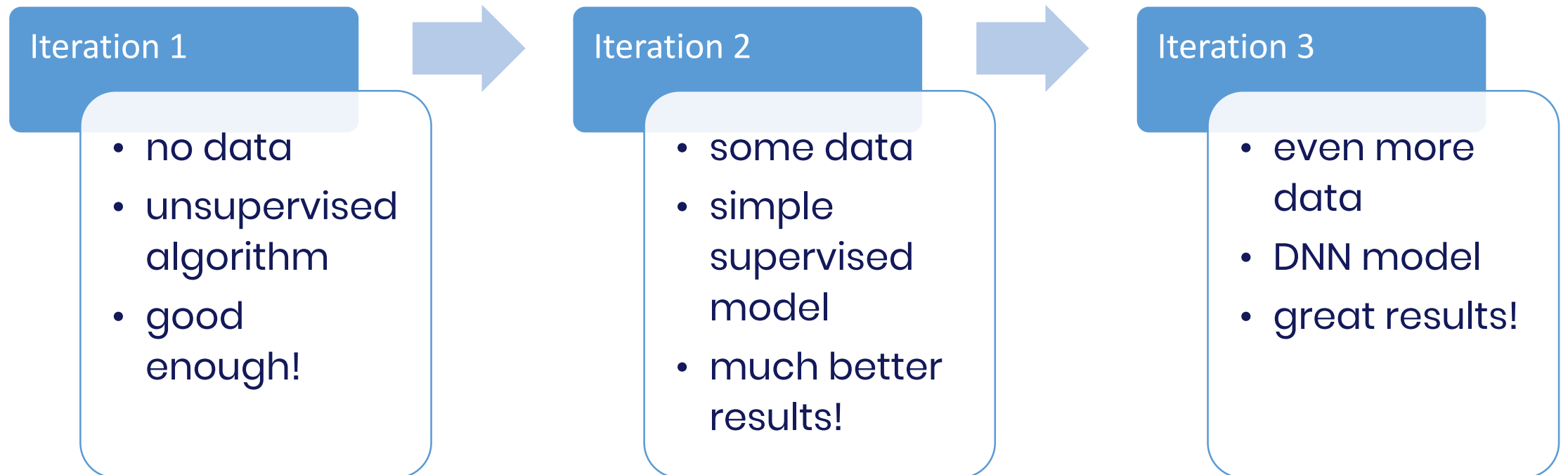


# Model development

- Challenges
  - Inconsistent annotations
  - Huge variation in accuracy on different data samples
- Solution
  - “debug” annotations using model assistance
- Result
  - Quickly got excellent results without tuning model.



# Example 2: snack detection



# Take aways

- Get started with something simple, setup a baseline.
- Unsupervised approach – easy to start, but has limitations.
- Supervised approach is worth the effort.
- Data is more important, than the model.
- Model-assisted approach helps to quickly identify bugs in data.

# ML Project Lifecycle



- Monitoring
- Maintaining

## Challenges

- ML issues
  - Data drift
  - Unexpected inputs
- Software engineering issues

# Data Drift

- Training set:
  - podcasts, news
- Real inputs:
  - zoom meetings, webinars

# Monitoring

- Model inputs and outputs
  - Trends
  - Outliers
- Client feedback
- System health

# Monitor Model Inputs / Outputs

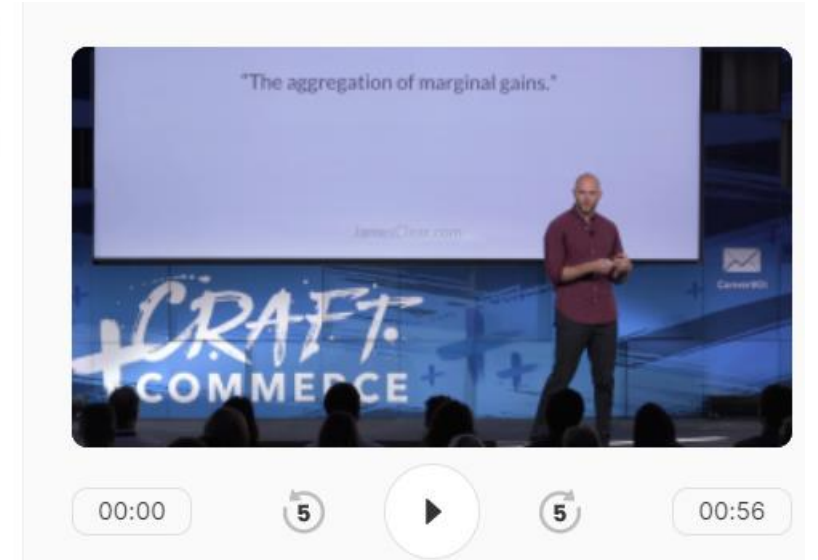
- Manual analysis
  - e.g. examine 10 random files a day
- Track metrics
  - Inputs
    - File type
    - File duration
    - Number of speakers
    - Word rate
  - Model Outputs
    - Number of chapters predicted
    - Smallest/largest chapter duration

# Customer Feedback

◀ Trimming audio ⓘ

Add/remove segments

last five. Now I've got two British cyclists, but it was at the Olympics in London in 2012 when this kind of strategy really came to fruition. They won 70% of the gold medals available. And so this idea that small improvements, tiny habits, little choices are not just a cherry on top of our performance, not just like a nice thing to have, but actually can be the key that unlocks significant success. That's an idea that I want us to carry with us as we go through the rest of this presentation.



## Implicit feedback

- Editing: snack boundaries, transcript.
- User exports a snack on YouTube.

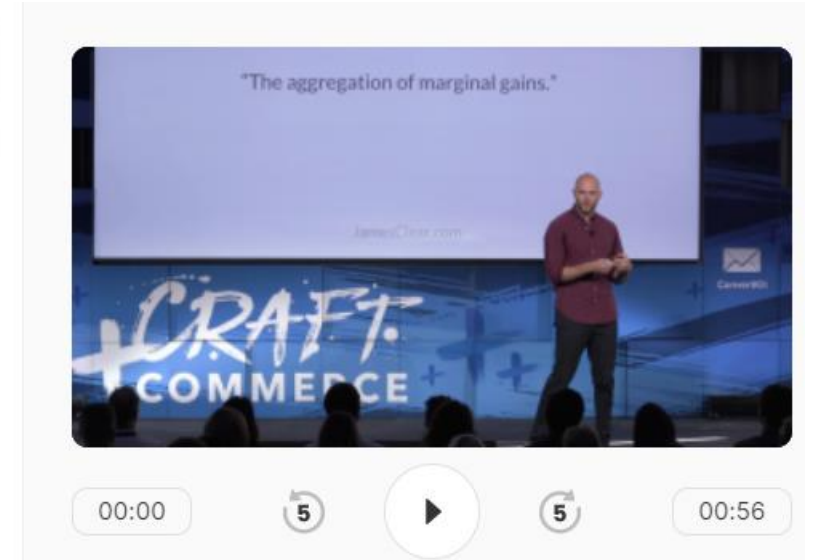


# Customer Feedback

◀ Trimming audio ⓘ

Add/remove segments

last five. Now I've got two British cyclists, but it was at the Olympics in London in 2012 when this kind of strategy really came to fruition. They won 70% of the gold medals available. And so this idea that small improvements, tiny habits, little choices are not just a cherry on top of our performance, not just like a nice thing to have, but actually can be the key that unlocks significant success. That's an idea that I want us to carry with us as we go through the rest of this presentation.



## Implicit feedback

- Editing: snack boundaries, transcript.
- User exports a snack on YouTube.

## Explicit feedback

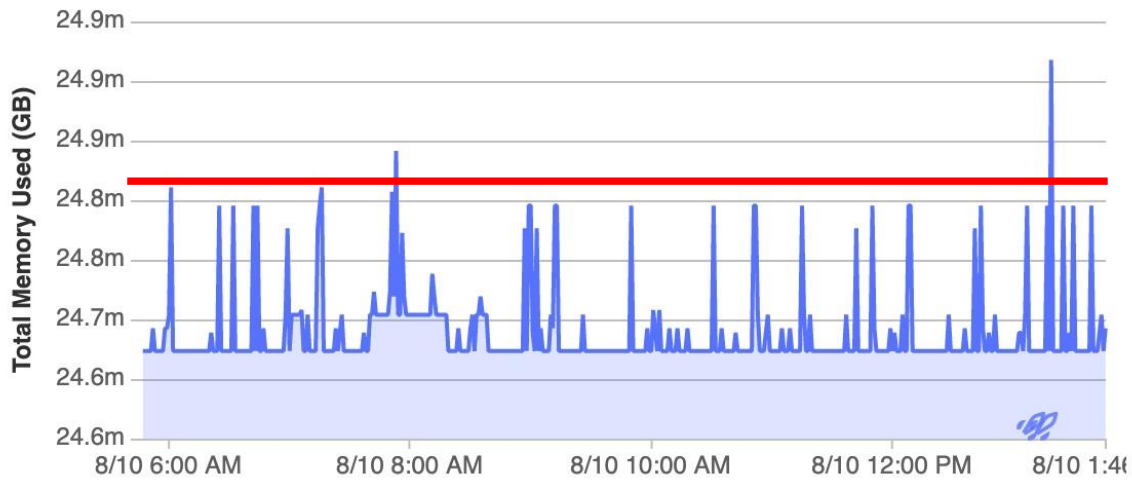
- Thumbs up/down
- Surveys/interviews

# Software/hardware metrics

- Server load
- Resource utilization (cpu, memory, disk, etc.)
- Processing time
  - How long it takes to process one hour of content
- Number of failed files per day.

# Dashboards

Memory Usage (Total)

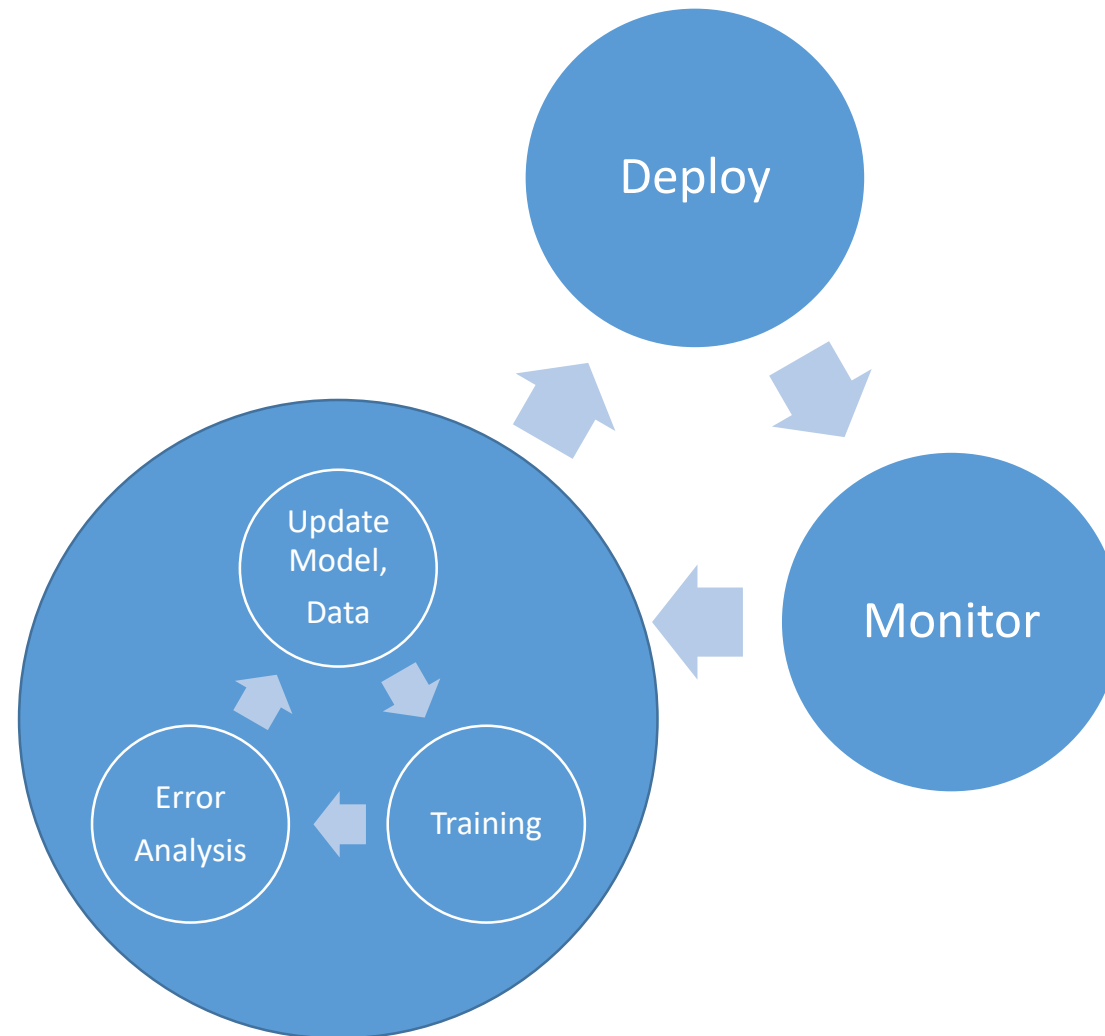


Core Usage (Total)



# Iterative Deployment

- Manual
- Automatic



# Roles in ML

- Researcher
- Data Scientist
- ML Engineer
- Data Engineer
- MLOps Engineer

**Thank you!**

Questions?