

How We Recommend Alternatives

Dr. Robert Kübler

About me



Dr. Robert Kübler
Data Scientist
Recommender Systems

A Nightmare!



METRO Onlineshop Angebote & Märkte Lieferservice Gastronomie

Im gesamten Sortiment suchen

Angebote Food Non-Food Produktwelten

Startseite > Food > Käse

WERBUNG (von 11-08-2022 bis 24-08-2022)

METRO Chef Gouda - 500 g Packung

953846



1 Packung

Zur Li

Verfü

Verfü

Produ

Lagerun

Tiefgek

Nettoin

500 g

Produkt

20

Regulie

Gouda

Von-Bis

25

Artikel

400025

Marke

METRO

Beschreibung Zutaten und Ernährungshinweise Eigenschaften

A Nightmare!



METRO Onlineshop Angebote & Märkte Lieferservice Gastronomie

Im gesamten Sortiment suchen


Angebote Food Non-Food Produktwelten

Startseite > Food > Käse

WERBUNG (von 11-08-2022 bis 24-08-2022)

METRO Chef Gouda - 500 g Packung

953846



You are too late, everything is gone.

Deal with it, loser.

Beschreibung Zutaten und Ernährungshinweise Eigenschaften

1 Packung
Zur Li
Verfü
Verfü
Produ
Lagerun
Tiefgek
Nettoin
500 g
Produk
20
Regulie
Gouda
Von-Bis
25
Artike
400025
Marke
METRO

Much better



METRO Onlineshop Angebote & Märkte Lieferservice Gastronomie

Im gesamten Sortiment suchen

Angebote Food Non-Food Produktwelten



Startseite > Food > Käse

WERBUNG (von 11-08-2022 bis 24-08-2022)

METRO Chef Gouda - 500 g Packung

953846

Sorry, the article you searched for is currently not available.
But look at these awesome replacements instead. 😊



Beschreibung Zutaten und Ernährungshinweise Eigenschaften

1 Packung
Zur Liste
Verfügbar
Produkt
Lagerung
Tiefgefroren
Nettoinhalt
500 g
Produkt
20
Regulierung
Gouda
Von-Bis
25
Artikelnummer
400025
Marke
METRO

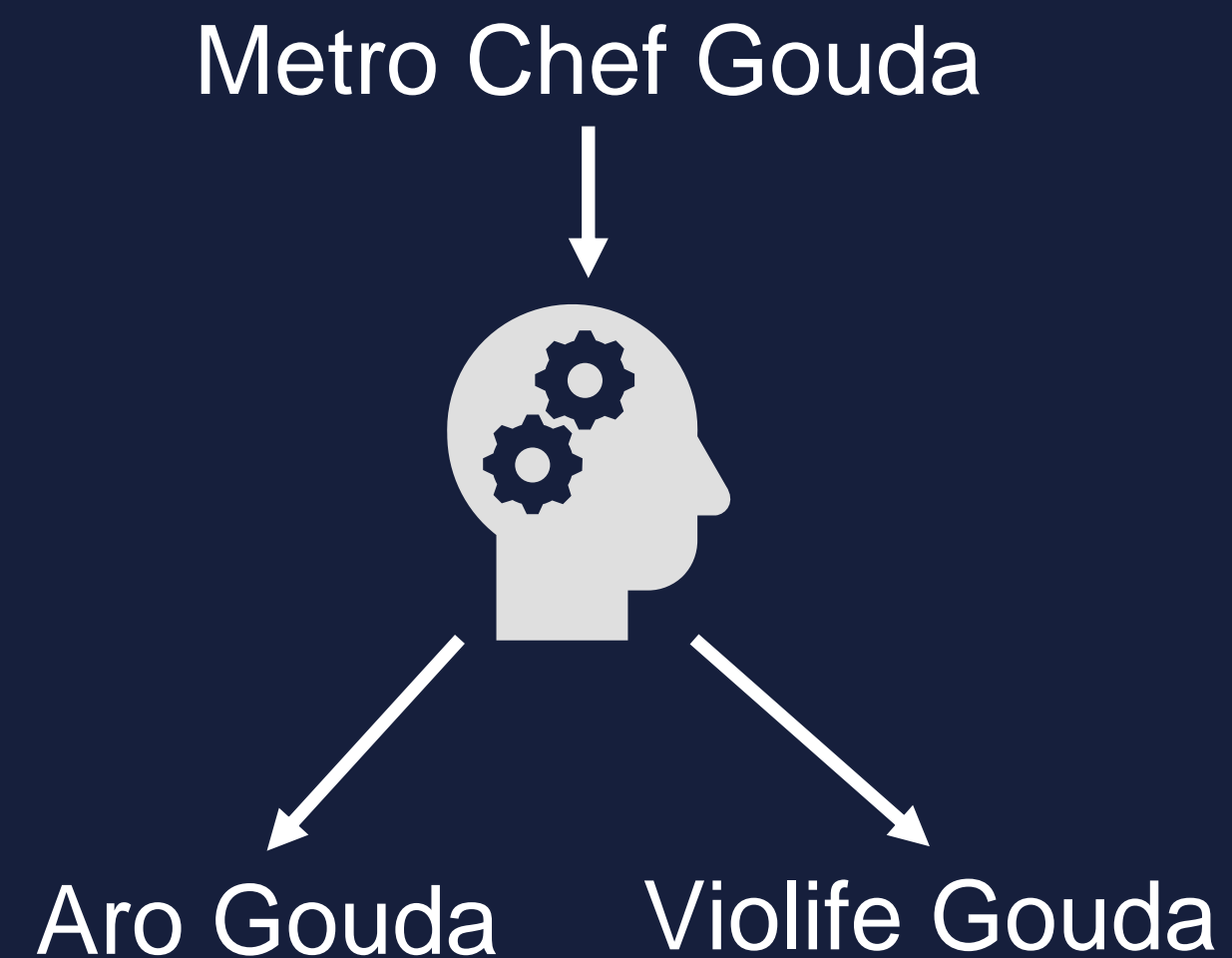
High-Level Overview



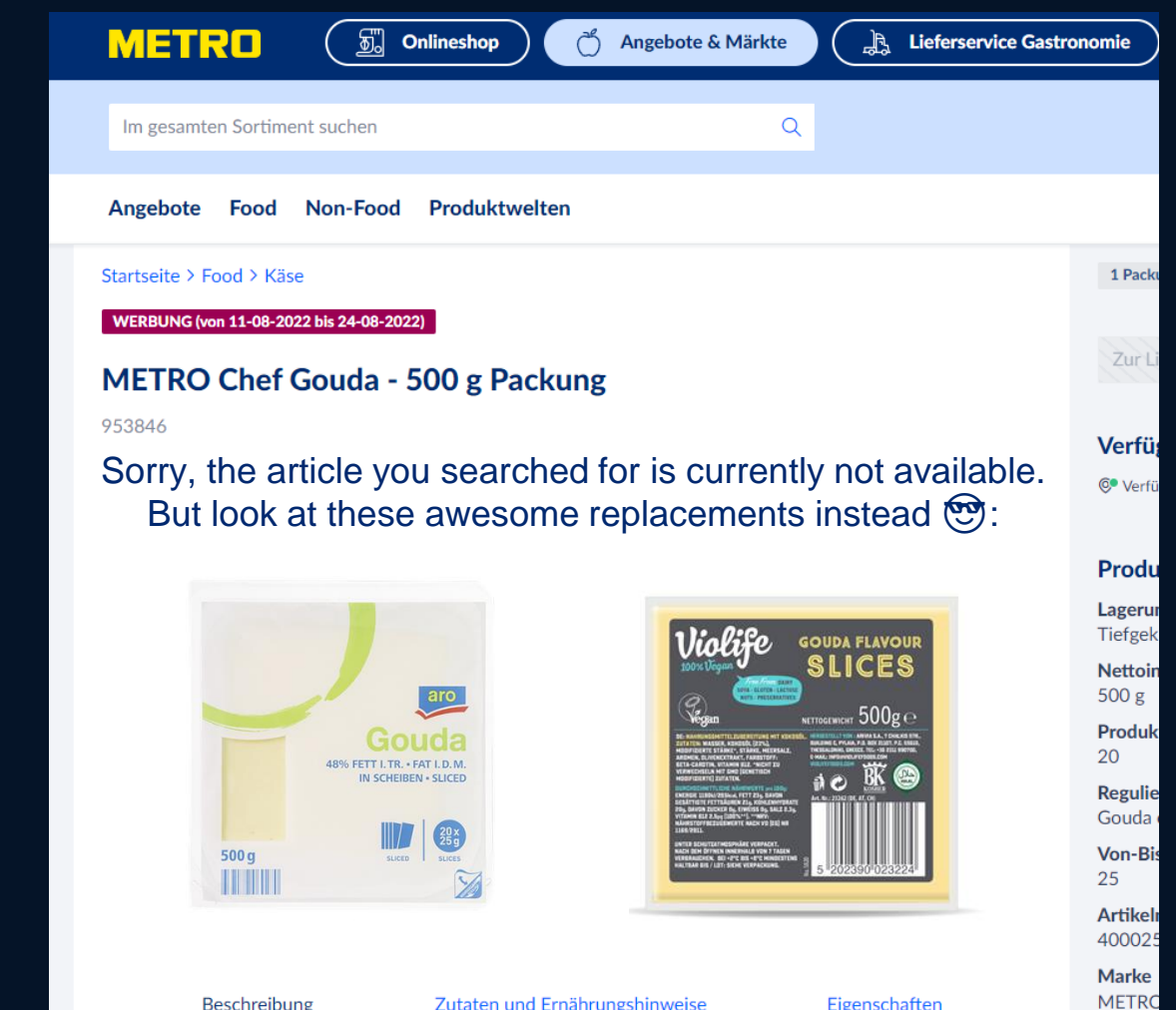
1 Get data

Missing Article	Replacement
Metro Chef Gouda	Aro Gouda
Metro Chef Gouda	Violife Gouda
Cola	Cola Light
Cola	Pepsi

2 Train model



3 Serve



Getting the Data

Life is Not Easy



Missing Article	Replacement
Metro Chef Gouda	Aro Gouda
Metro Chef Gouda	Violife Gouda
Cola	Cola Light
Cola	Pepsi

Data From Implicit Feedback



Raw data: Online User Sessions!

Cola 0.33 l

Cola 0.5 l

Pepsi

Cola light

Pepsi light

Butter

Main assumption:

Alternative products are neighbors.



Data From Implicit Feedback



Cola 0.33 l

Cola 0.5 l

Pepsi

Cola light

Pepsi light

Butter



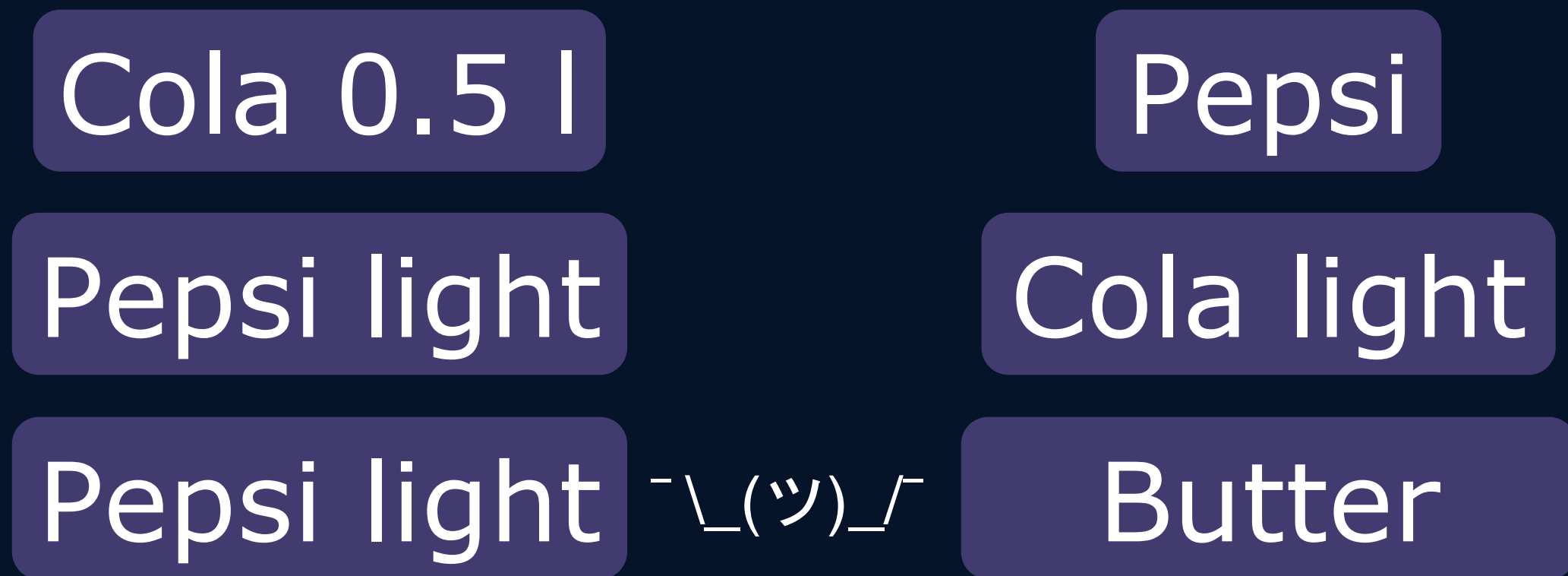
Creating the Dataset



Not necessarily direct neighbors!

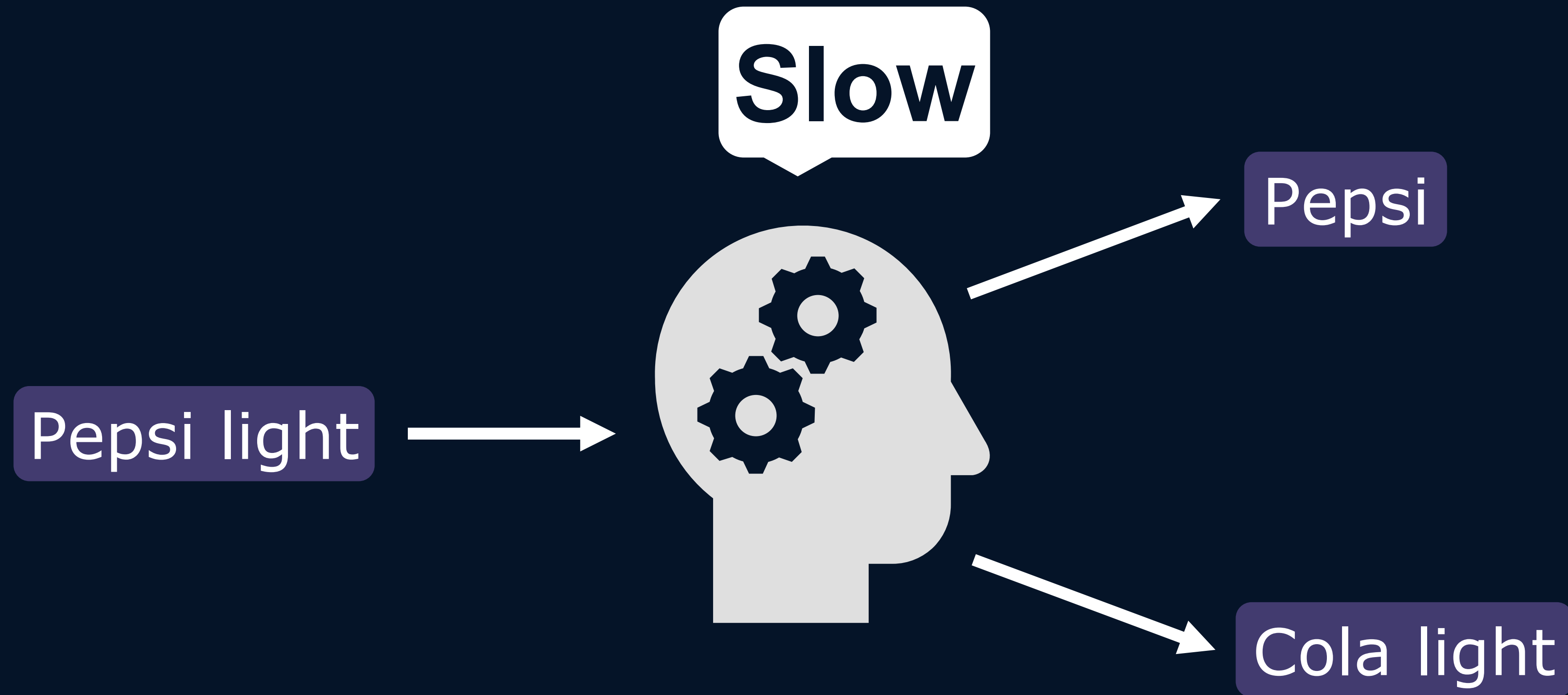
Missing Article

Replacement



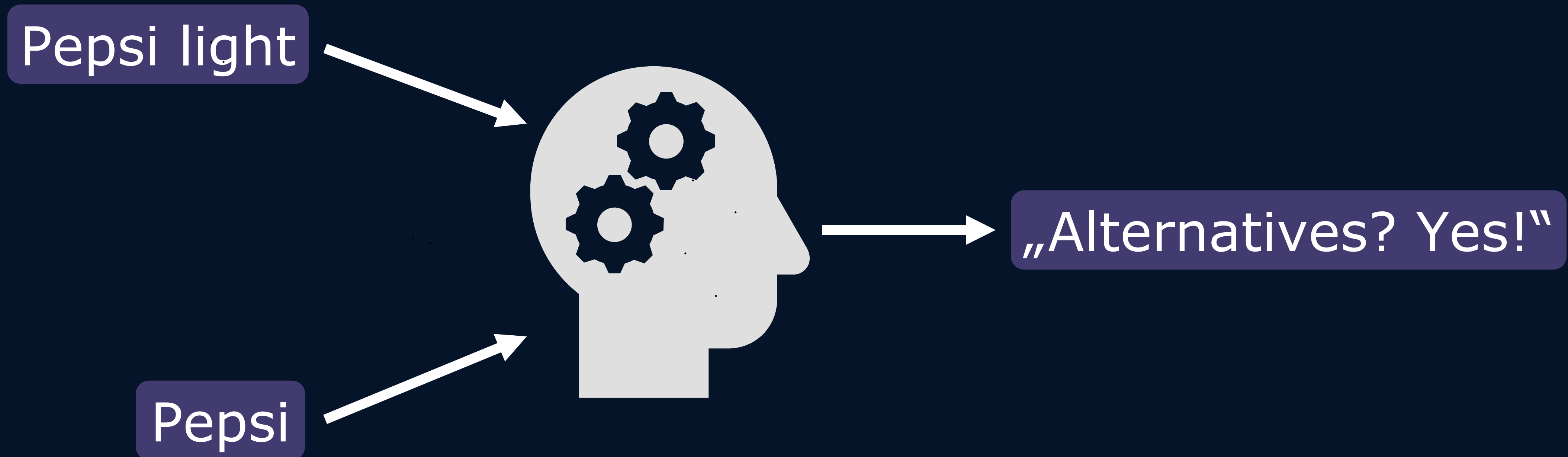
Training the Model

Direct Way

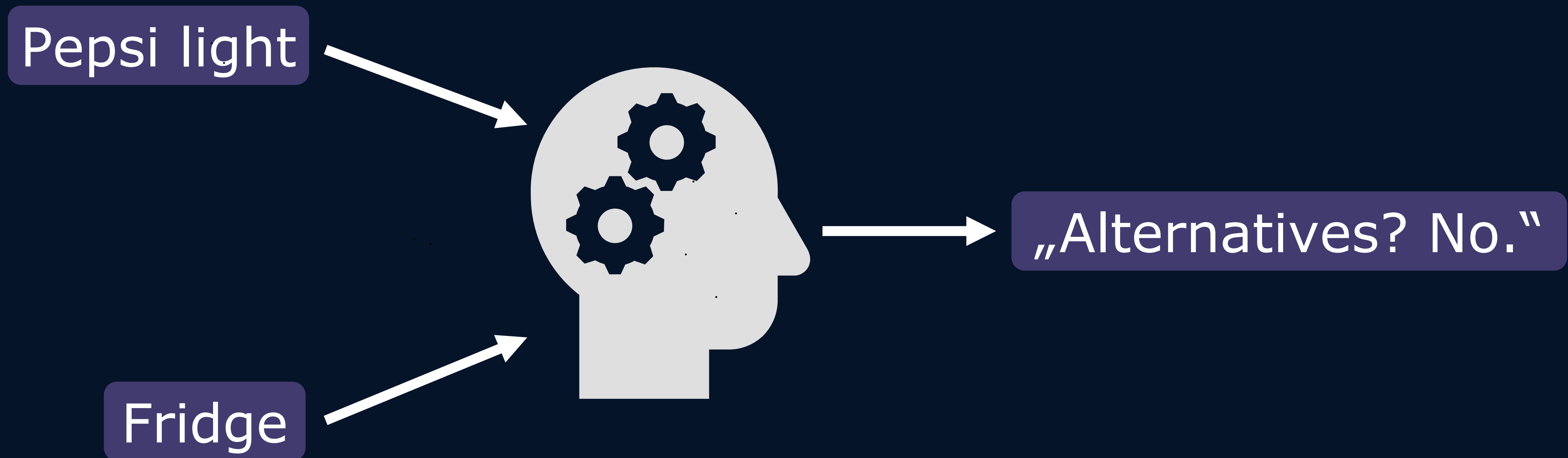


Multiclass-multioutput classification

Efficient Way



Efficient Way



Binary classification

~~Only a Single Class~~ Two Classes!



Cola 0.33 l

Cola 0.5 l

Pepsi

Cola light

Pepsi light

Butter

Missing Article

Replacement

Alternatives?

Cola 0.5 l

Pepsi

Yes

Pepsi light

Cola light

Yes

Pepsi light

Butter

Yes

Cola 0.5 l

Fridge

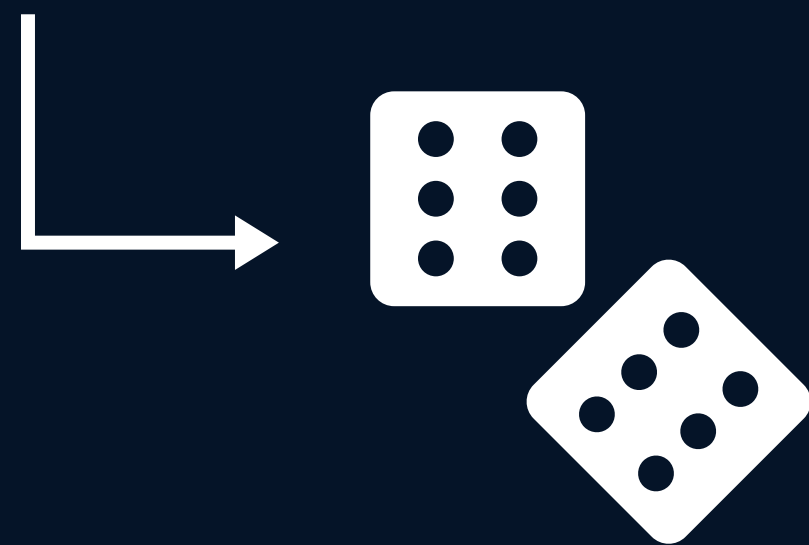
No

Butter

iPhone

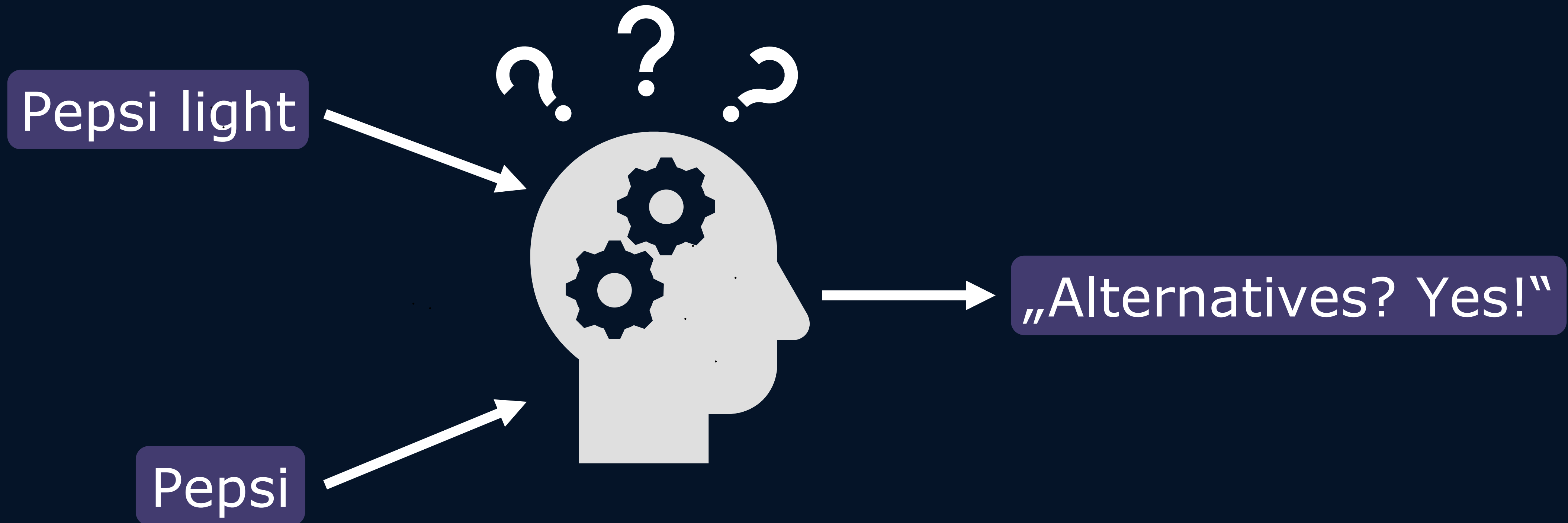
No

„Negative Sampling“

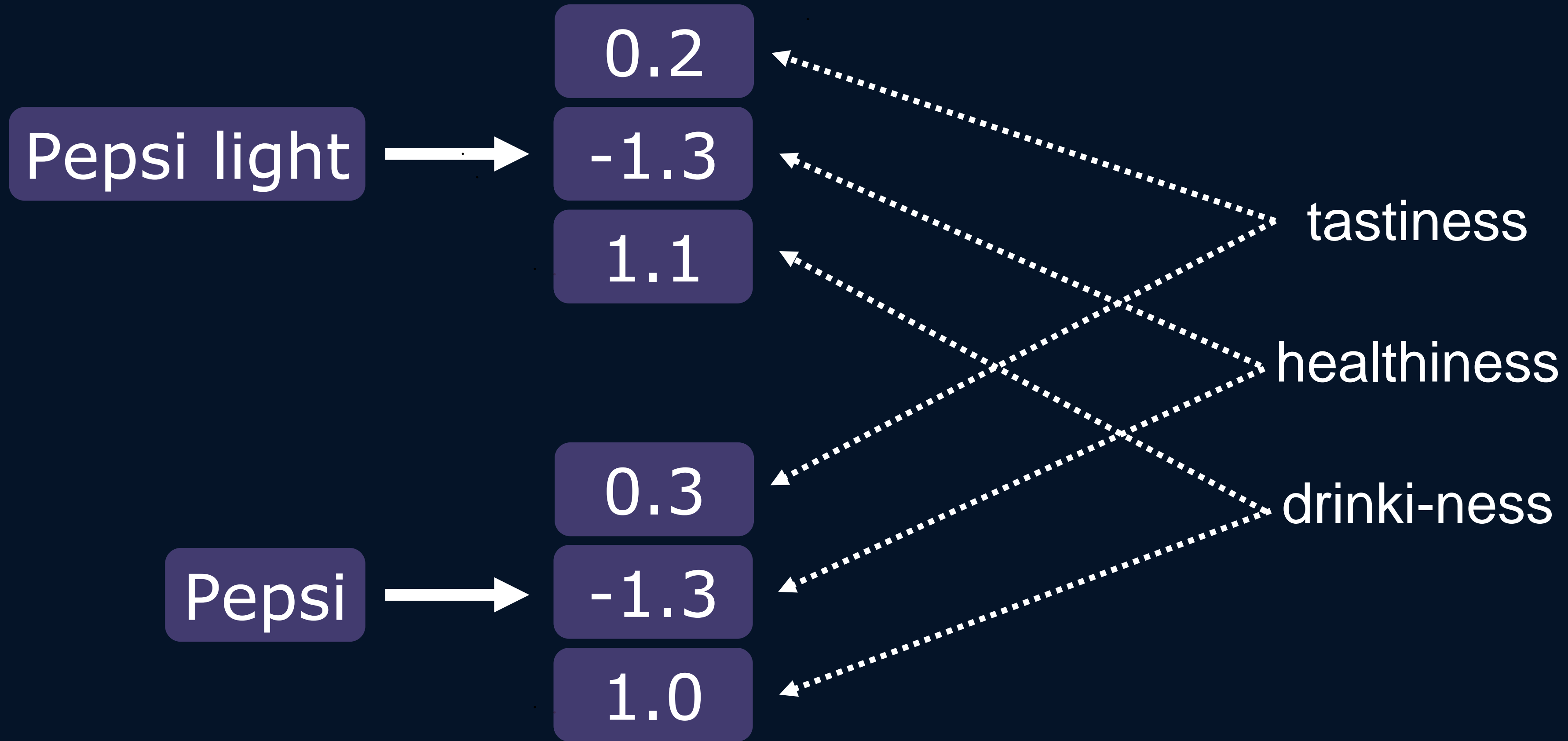


```
def model(art_1, art_2):  
    return "Yes"
```

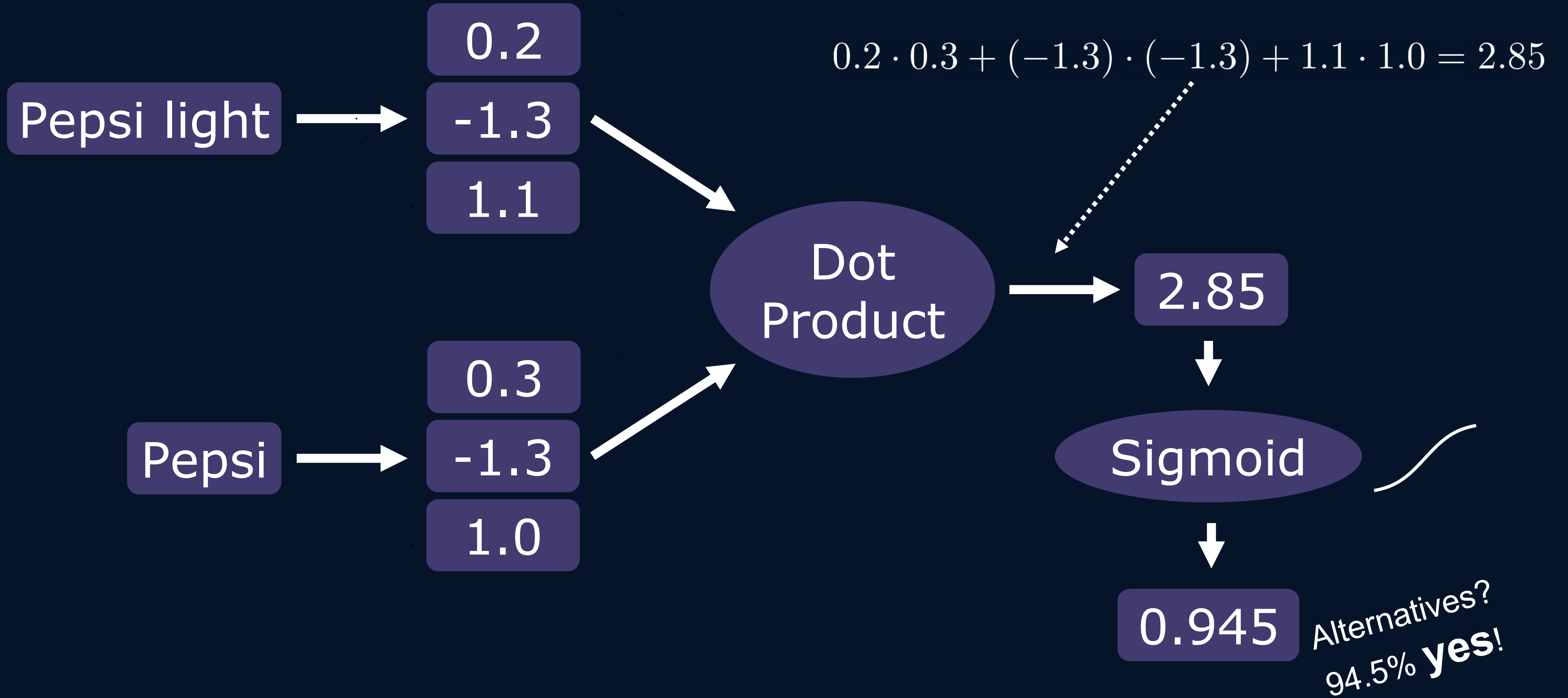
Turning Articles Into Numbers



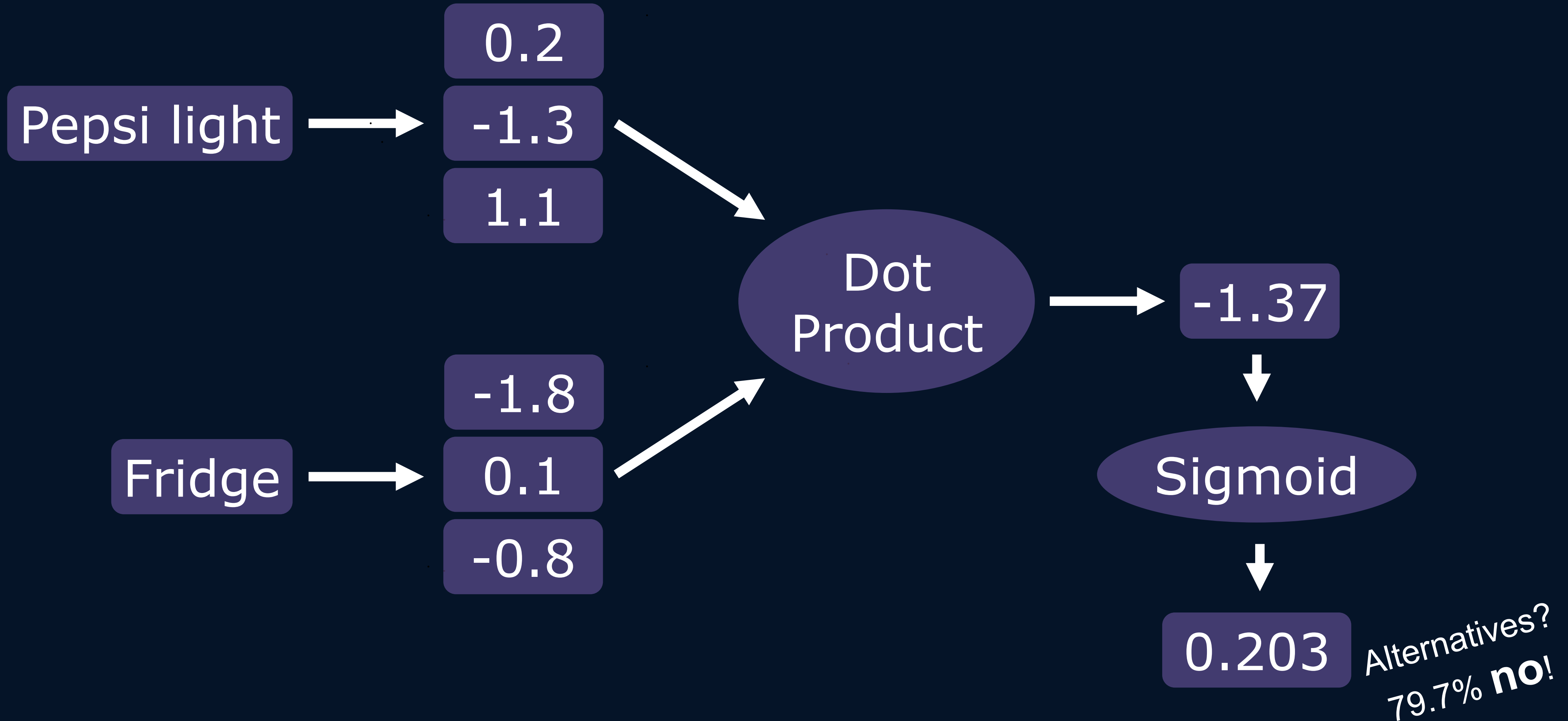
Turning Articles Into Numbers



Turning Articles Into Numbers



Turning Articles Into Numbers



Ok, great, we have trained a model.

But how can we make it useful?

Using the Model



Throw away the model! Only use the embeddings.



Great property: Alternatives have similar embeddings!

About the Embeddings

- Easy to serve, easy to use
- Can be reused in different contexts
- Typically high-dimensional, but project into two or three dimensions for nice visuals



Cold Start Problem



Article not seen during training?
No embedding then. 🙈

Takeaway so far



1. If explicit data is unavailable, use **implicit data**.
2. If your model is too slow, make it **simpler**.
3. If you deal with many categories, use **embeddings**.

Looks Familiar?



Word2vec!

words = articles
sentences = online sessions



Gensim

„The fit-predict of NLP“



Tensorflow



PyTorch

Where to go from here?

Use Product (Meta) Data!



Product: Pepsi 1l
Category: Softdrink
Brand: Pepsi
Quantity Type: Bottle
Price: Premium



Use for training as well!

Meta-Prod2Vec*



Meta-Prod2Vec - Product Embeddings Using Side-Information for Recommendation

Flavian Vasile
Criteo
Paris
f.vasile@criteo.com

Elena Smirnova
Criteo
Paris
e.smirnova@criteo.com

Alexis Conneau *
Facebook AI Research
Paris
aconneau@fb.com

ABSTRACT
We propose Meta-Prod2Vec, a novel method to compute item similarities for recommendation that leverages existing item metadata. Such scenarios are frequently encountered in applications such as content recommendation, ad targeting and web search. Our method leverages past user interactions with items and their metadata to improve the quality of recommendations. It addresses the cold-start problem [5] and handling the cold-start problem [29].
In the last couple of years, a promising new class of neural probabilistic models that can generate user and product embeddings has emerged and has shown promising results. The new methods can scale to millions of items and show good improvements on the cold-start problem. In the context of



Properties



- Embeddings for articles
- Embeddings for metadata
- Regularization



Properties



- Embeddings for articles
- Embeddings for metadata
- Regularization

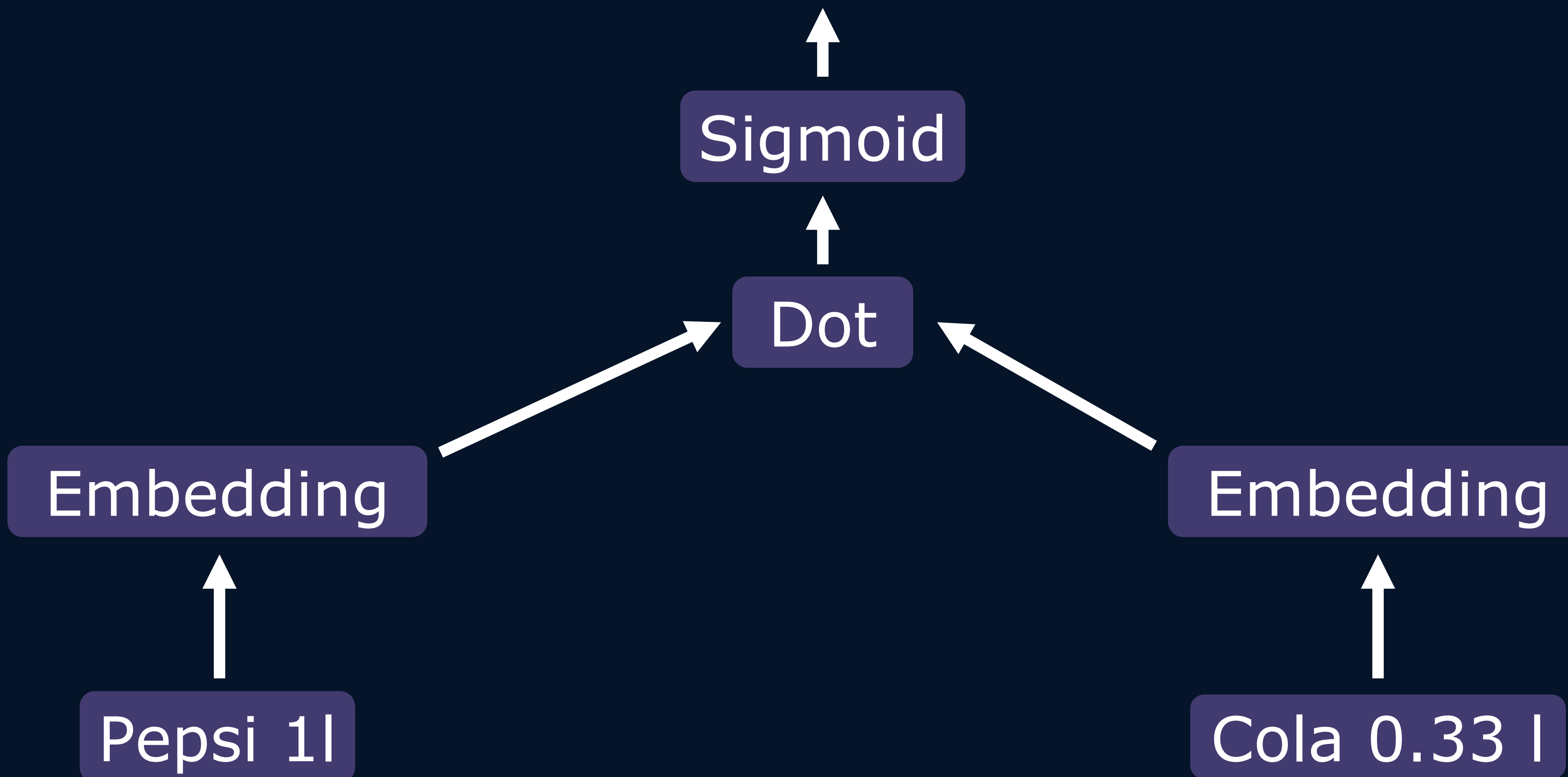


Article not seen during training?
Just use the metadata embeddings!

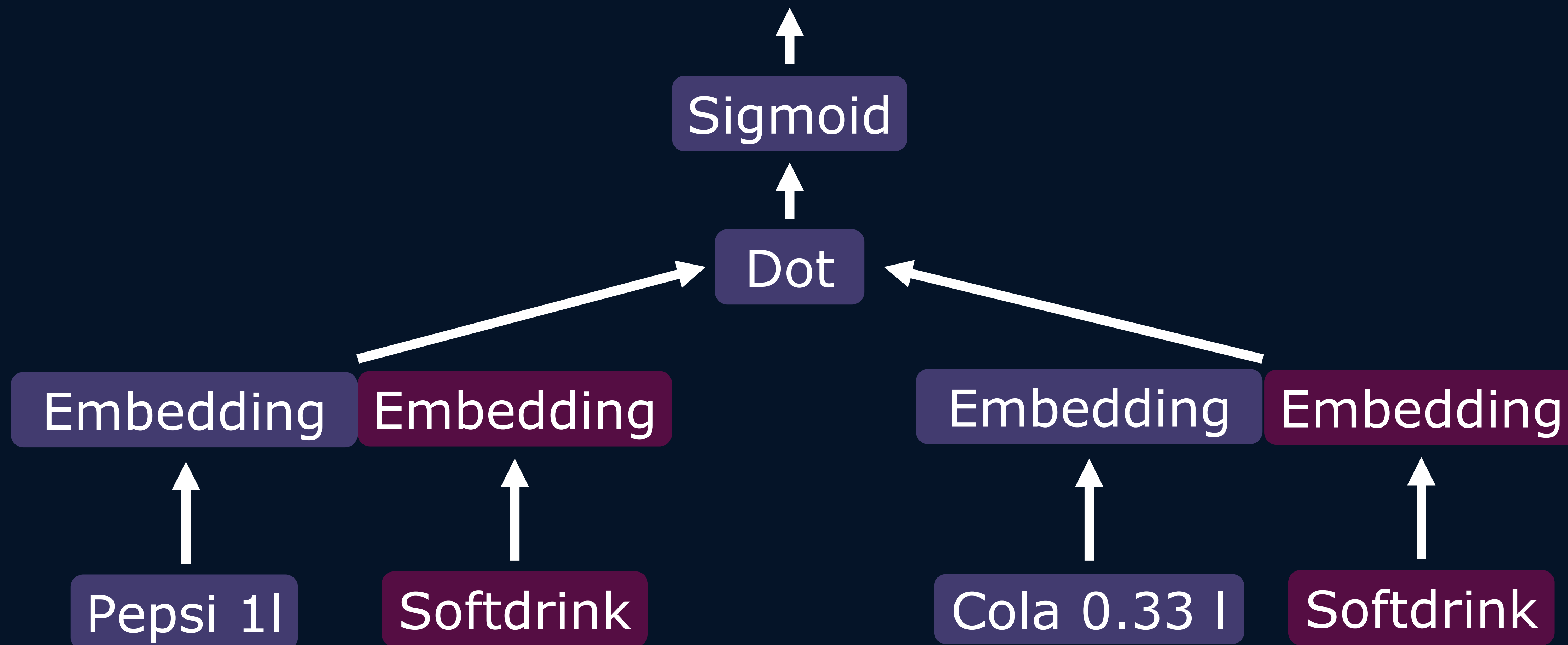


Try other architectures!

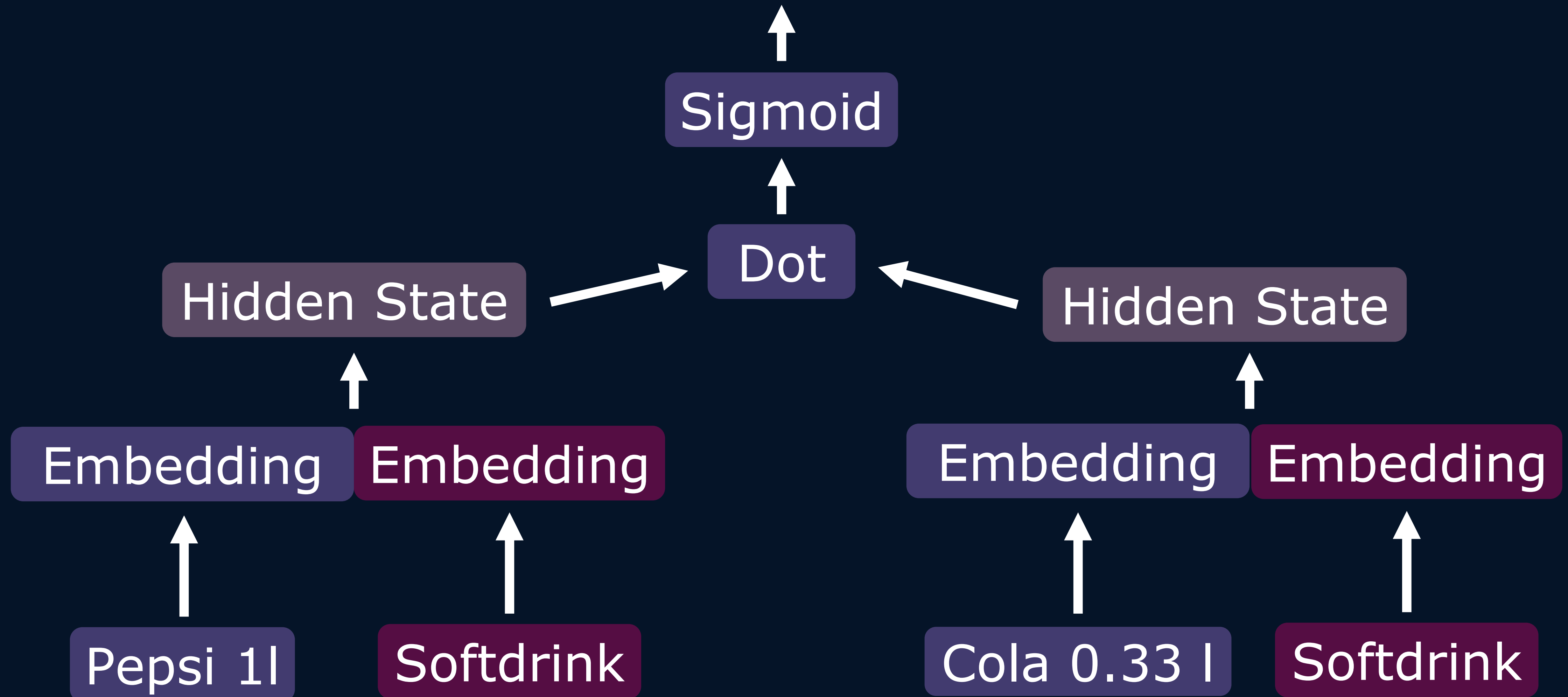
Two Towers Architecture



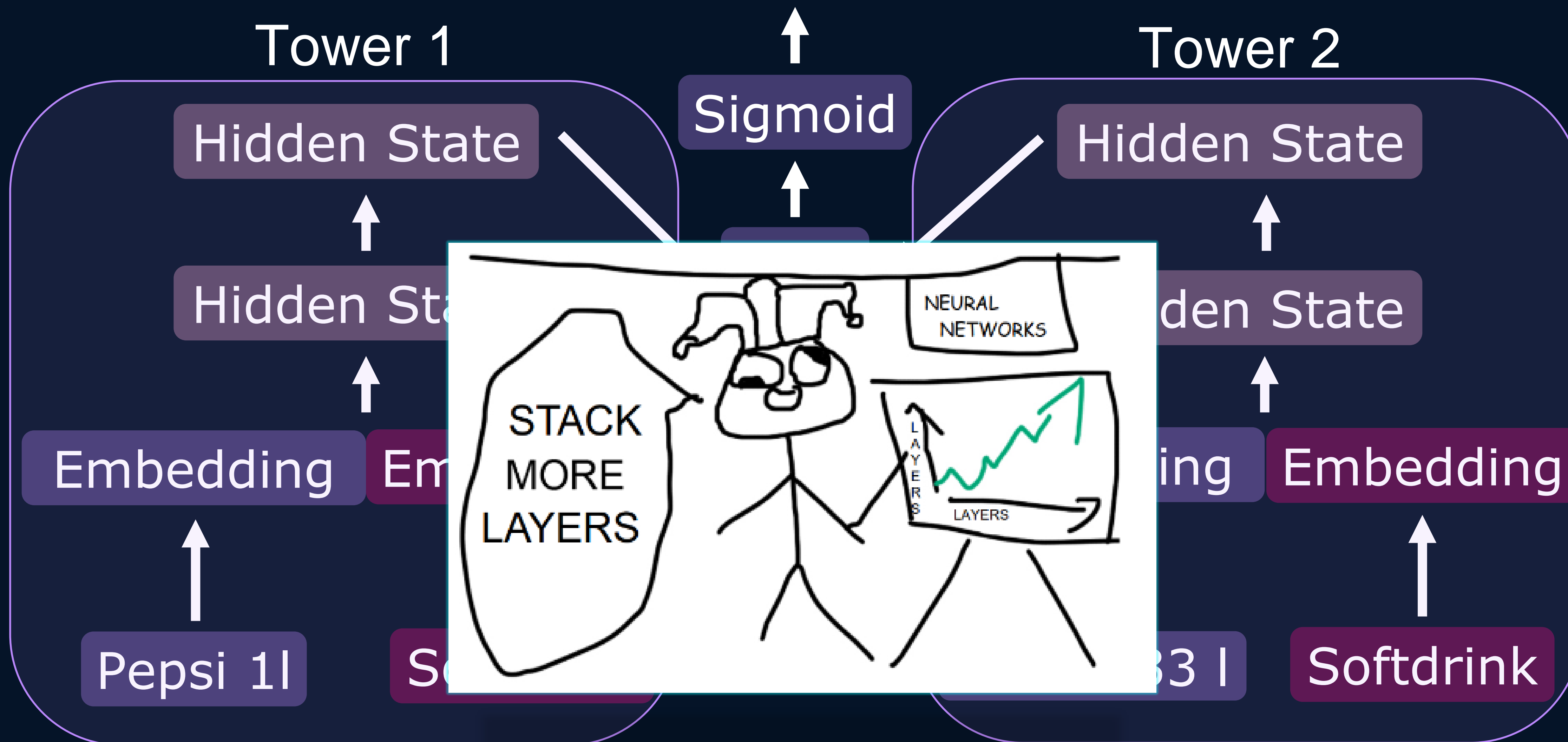
Two Towers Architecture



Two Towers Architecture



Two Towers Architecture



Contact Information

Company Website

metro.digital/hubs/data-science-hub

My Medium Profile
kuebler.ai

My LinkedIn
linkedin.com/in/robert-kuebler

Thanks :)