



THRIVE GLOBAL

AI Intention Classifier

SAMANTHA WELCH- TECHNOLOGY INTERN

ZAKERY CLARKE- DATA SCIENCE TEAM LEAD

DATA SCIENCE

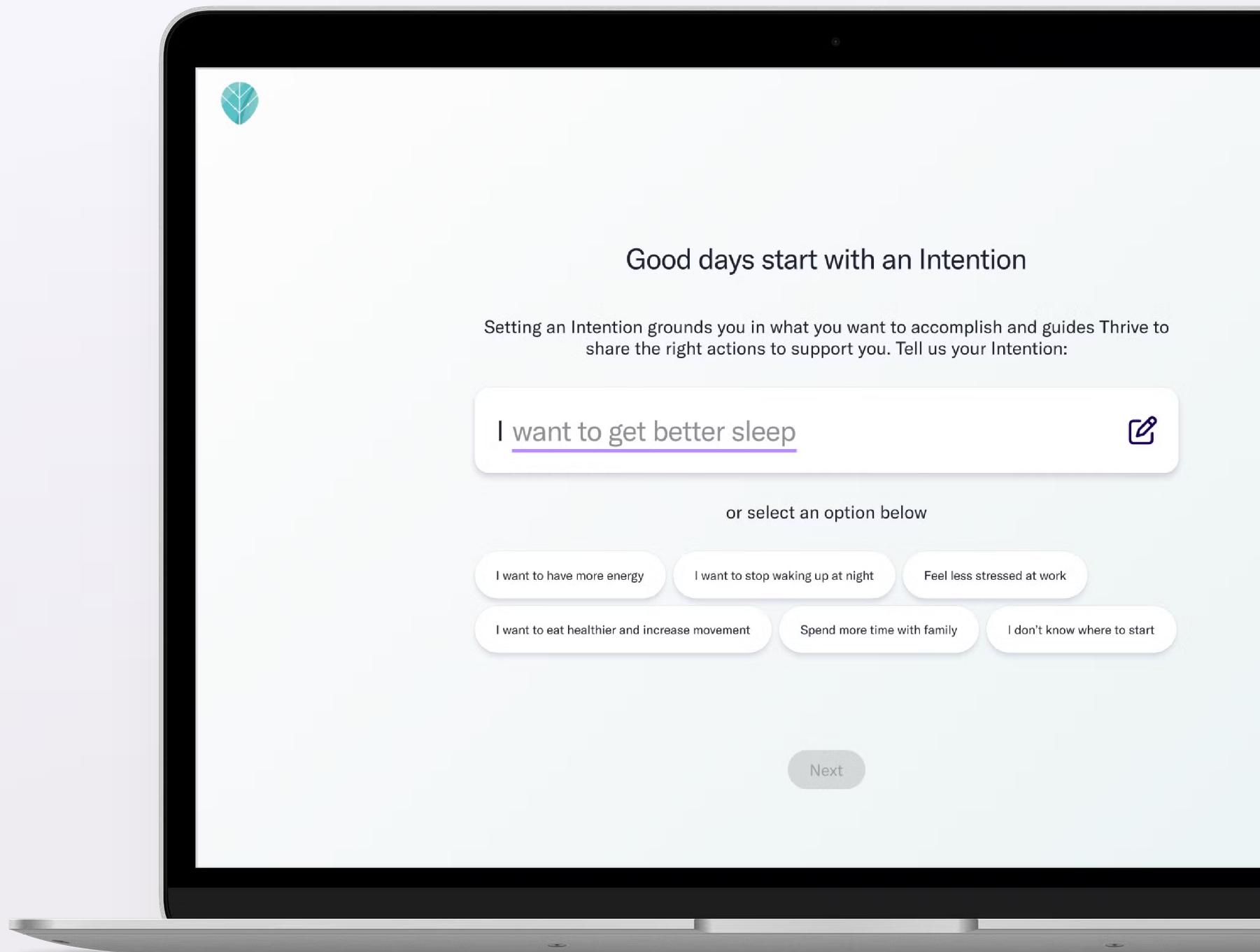


Project Overview

We are working in close collaboration with **#ask-microstep** to develop AI models that can personalize content through Intentions.

An **Intention** allows users to tell us where they want to focus on improving their well-being.

We have built an AI model that can map Intentions → Journey Areas so that we can personalize each user's experience



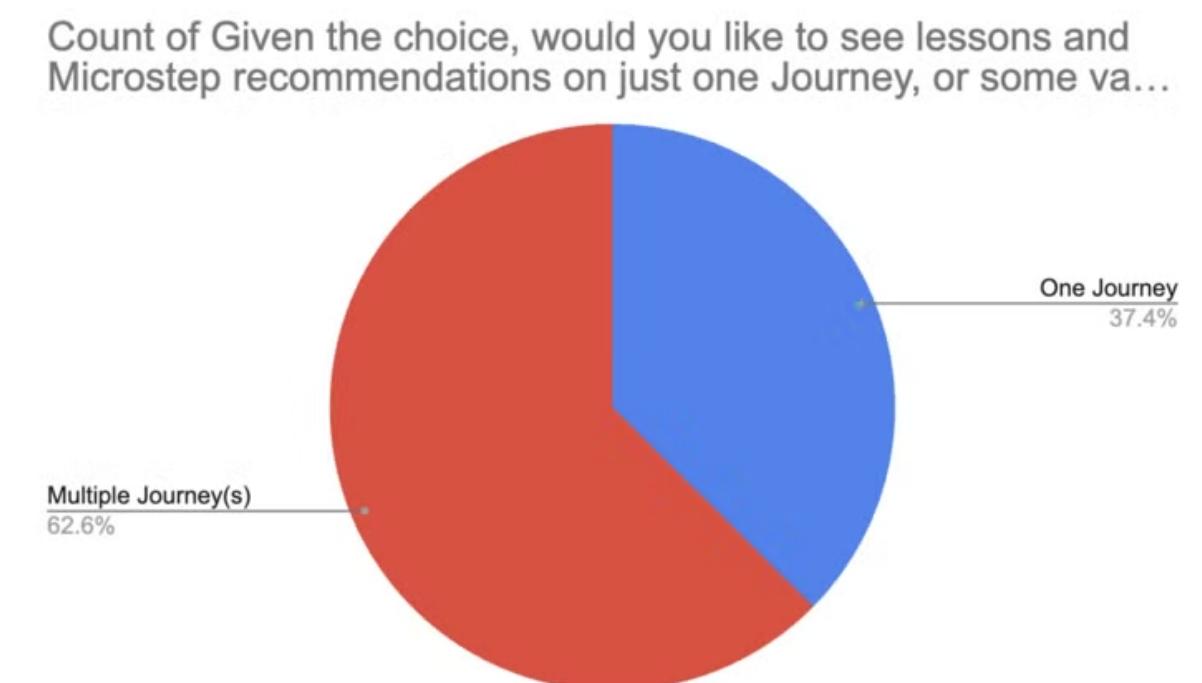
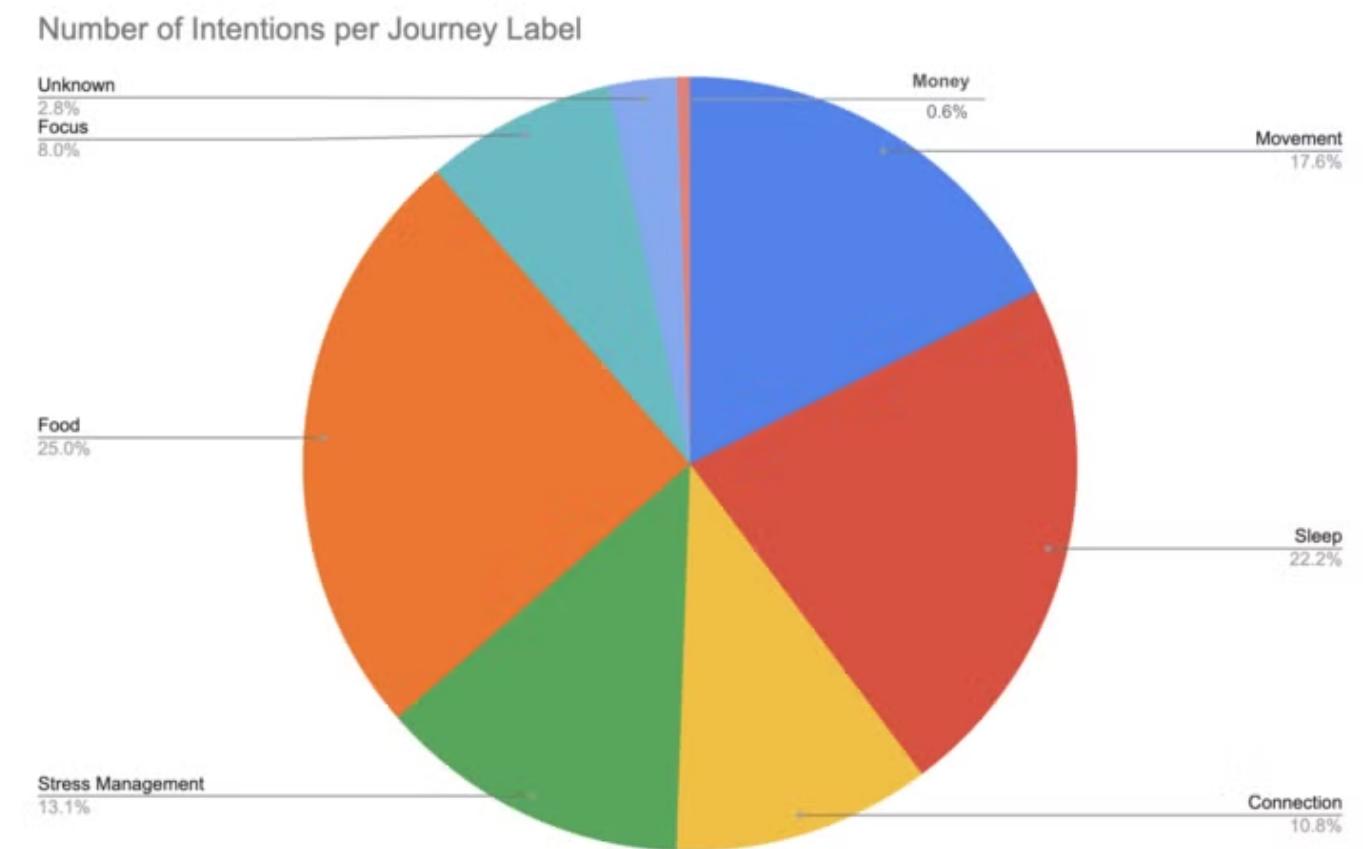
Survey Analysis



#ASK-DATA-SCIENCE

Intention Survey Results

- We received ~150 responses through an email survey with user's desired intentions
 - There are a variety of journeys in every intention, with the exception of Money
 - More people want multiple journeys vs. a single journey (63% vs. 37%)
 - 40% of users gave intentions that applied to multiple journey areas



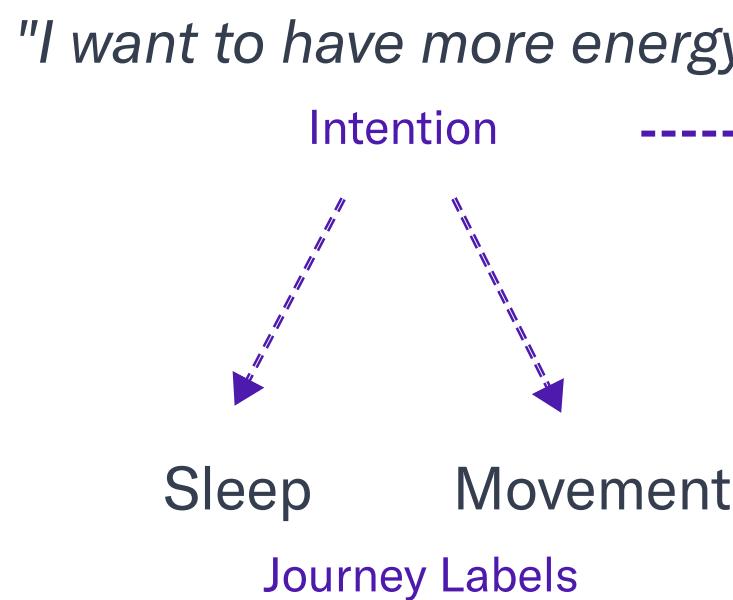
Intention Embeddings



#ASK-DATA-SCIENCE

How we map Intentions → Journeys using AI

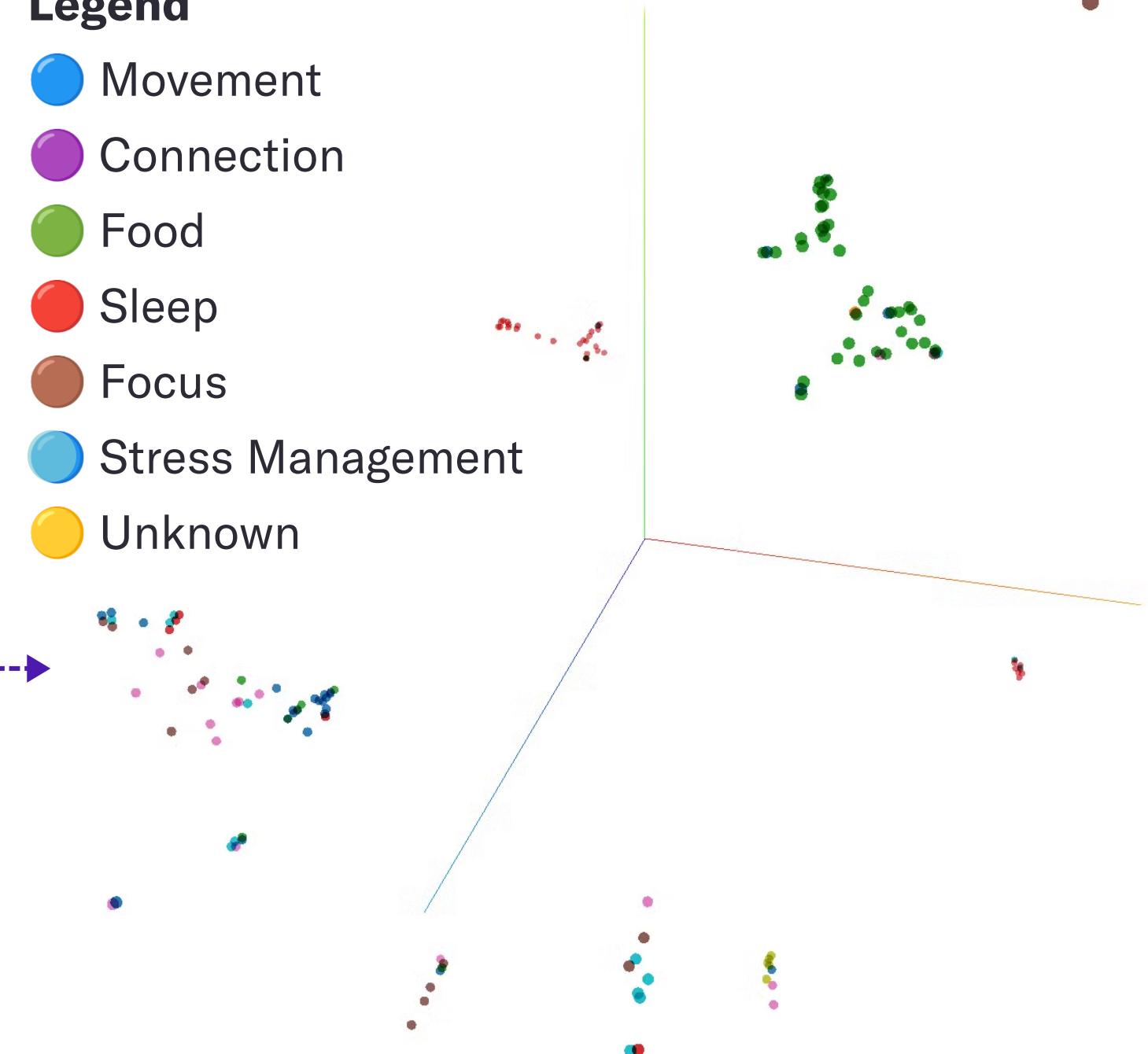
- Labeled journey(s) for each user intention
- Turn each intention into embeddings (numeric representation of words)
- Intentions that are neighbors in a cluster should have the same journey



Intention Embedding Clusters

Legend

- Movement (blue circle)
- Connection (purple circle)
- Food (green circle)
- Sleep (red circle)
- Focus (brown circle)
- Stress Management (cyan circle)
- Unknown (yellow circle)



Intention Embeddings



#ASK-DATA-SCIENCE

Comparison of Embeddings





Building an ML Model- Nearest Neighbors

#ASK-DATA-SCIENCE

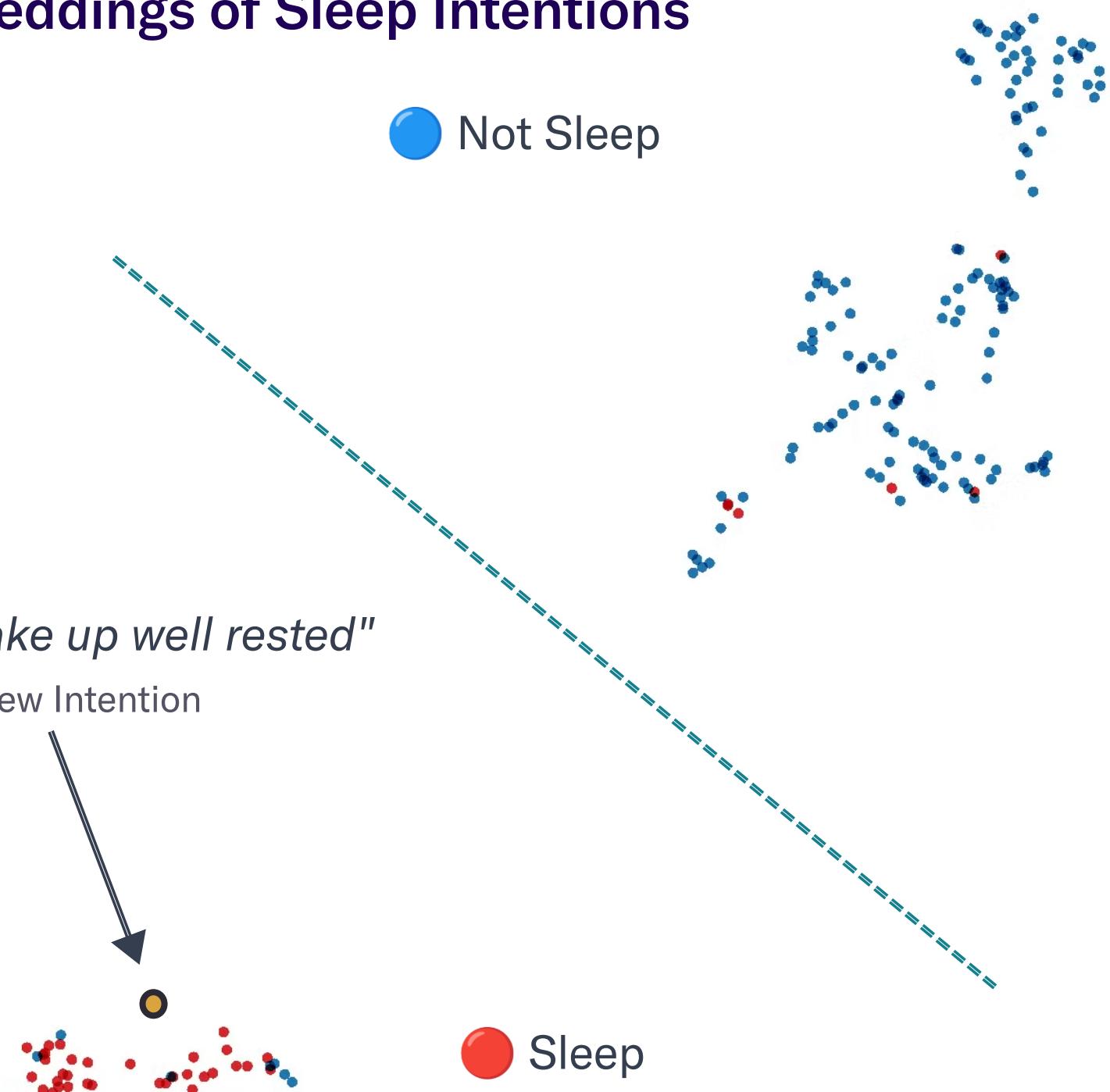
Model

K-Nearest Neighbors Classifier: classifies journey by using most common journey for nearby intentions

We created 1 model per Journey since an intention can map to multiple journeys → "Sleep" or "Not Sleep"

We experimented with different embedding models and different numbers of neighbors to improve the model accuracy

Embeddings of Sleep Intentions





Building an ML Model- Nearest Neighbors

#ASK-DATA-SCIENCE

Results

Sentence Transformers & OpenAI Embeddings performed the best, we chose to deploy Sentence Transformers because it is free and Open Source

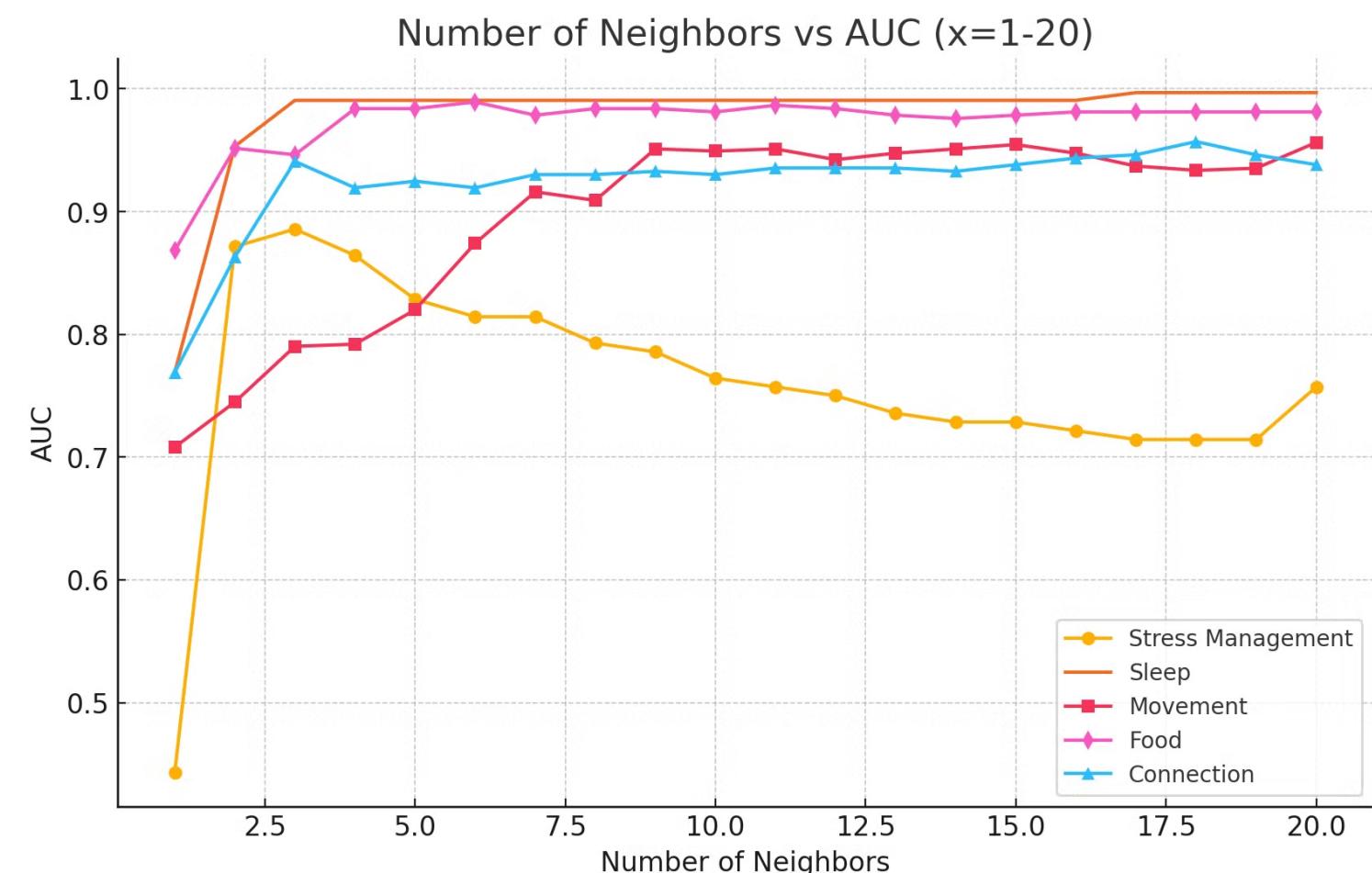
Number of Neighbors did not significantly impact performance as long as we selected values in the range of ~5-10 neighbors

AUC: How well the model classifies each intention

0.50: random chance

1.00: perfect accuracy

Embedding Model	Average AUC
TFIDF	0.78
Sentence Transformers ★	0.94
OpenAI	0.97



Building an ML Model- Nearest Neighbors



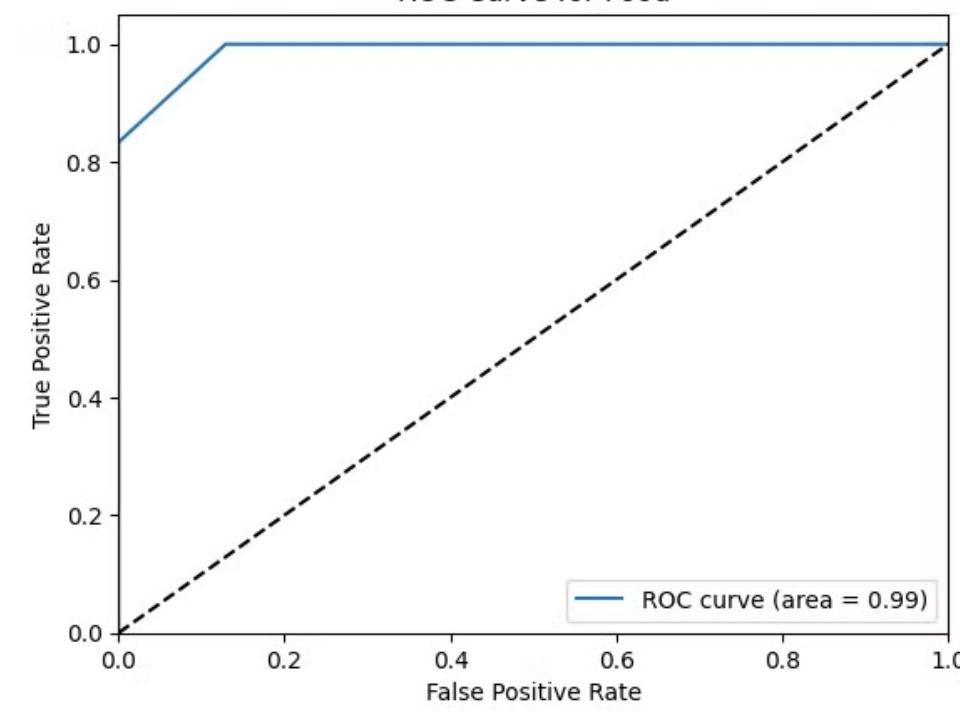
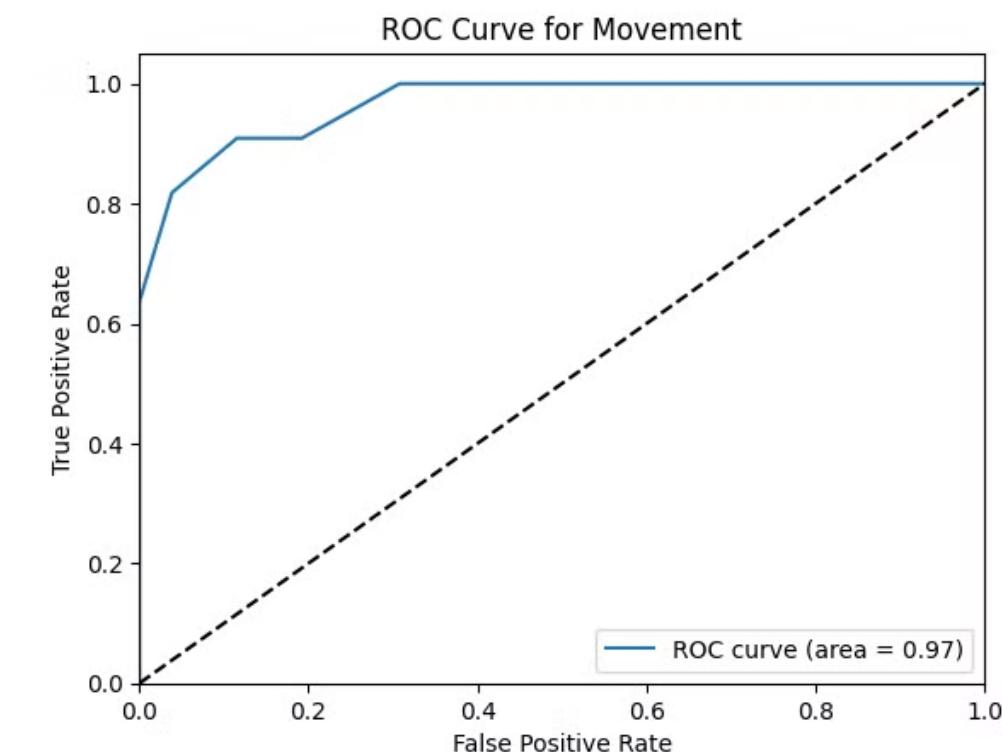
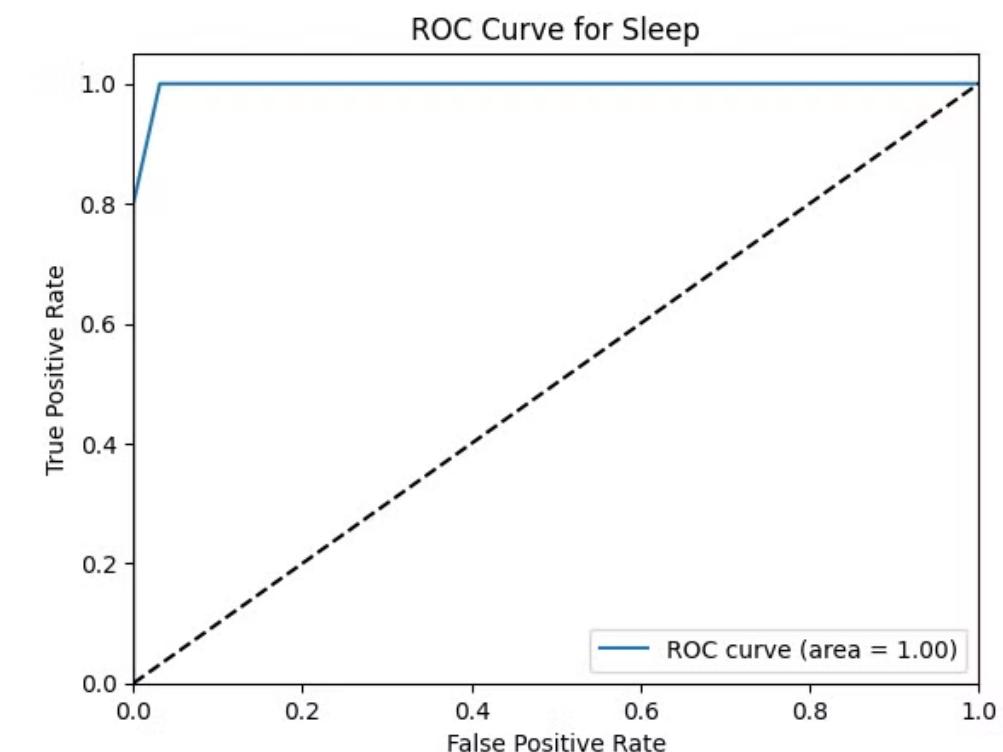
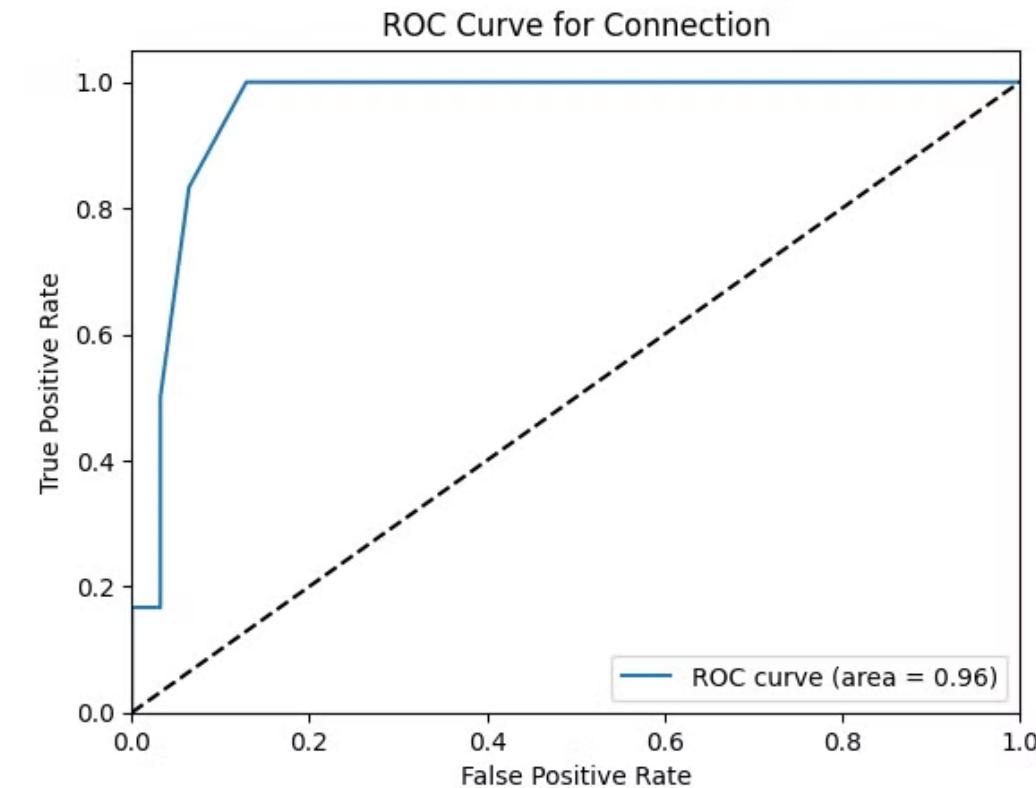
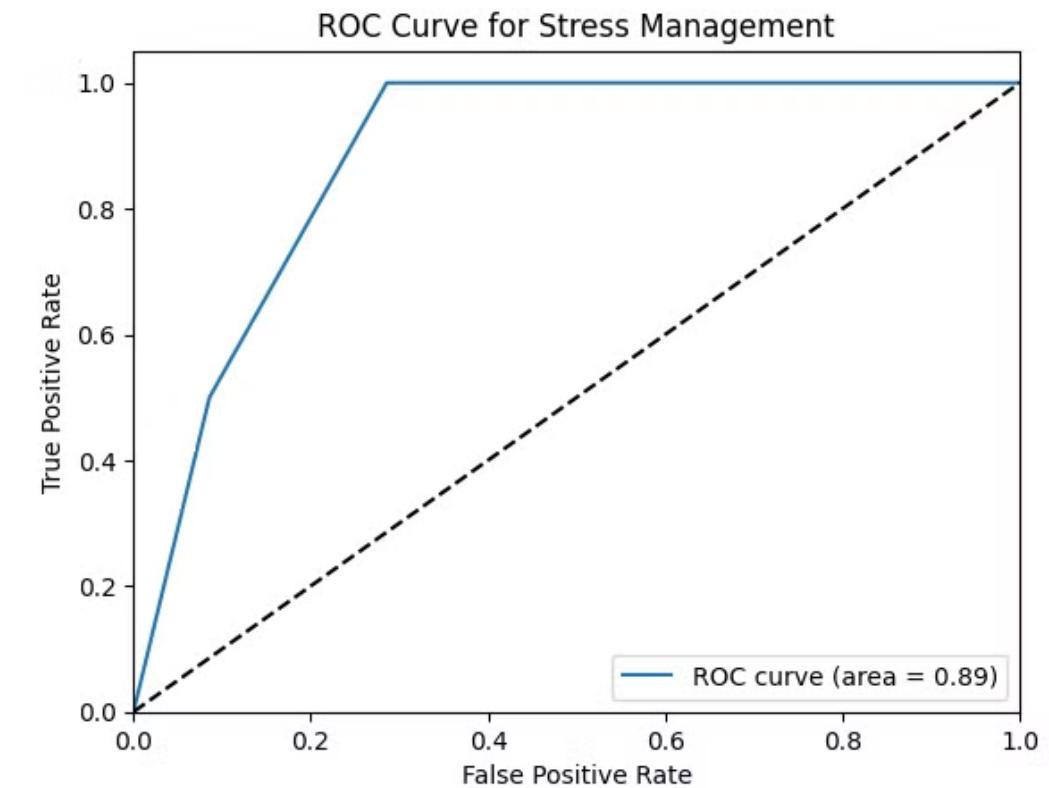
#ASK-DATA-SCIENCE

Best Model

SentenceTransformers ⭐

Number Neighbors = 5

Threshold = 0.3 FPR



Project Release

Launch in September 

#ask-data-science

#ask-microstep

“ Intentions will enable hyper-personalization of our product- ensuring each user receives a personalized Microstep plan that helps them best achieve behavior change.

Questions?



THRIVE GLOBAL