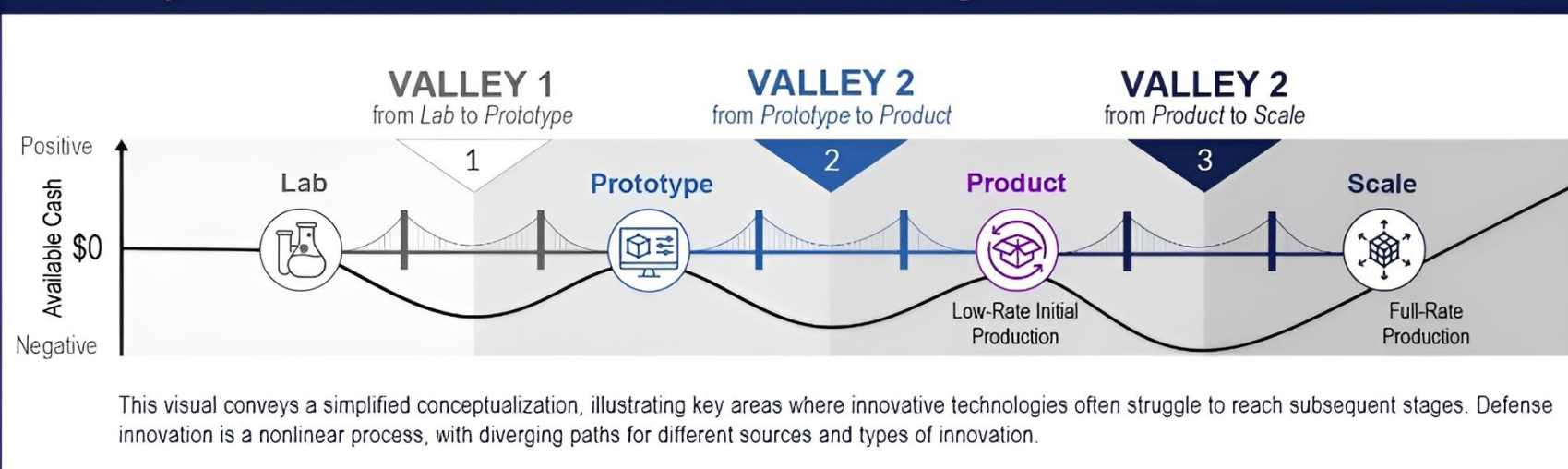


## PROBLEM STATEMENT

### Valleys of Death in Defense Innovation Acquisition



Technology commercialization has 4 main phases: feasibility (Lab), prototype, product, and full-scale production. Prior to full-scale production, that technology is often unprofitable. See references for image source.

Programs like SBIR are used to accelerate commercialization by small businesses.

**Goal:** Deliver business intelligence on strategic U.S. Government investments.

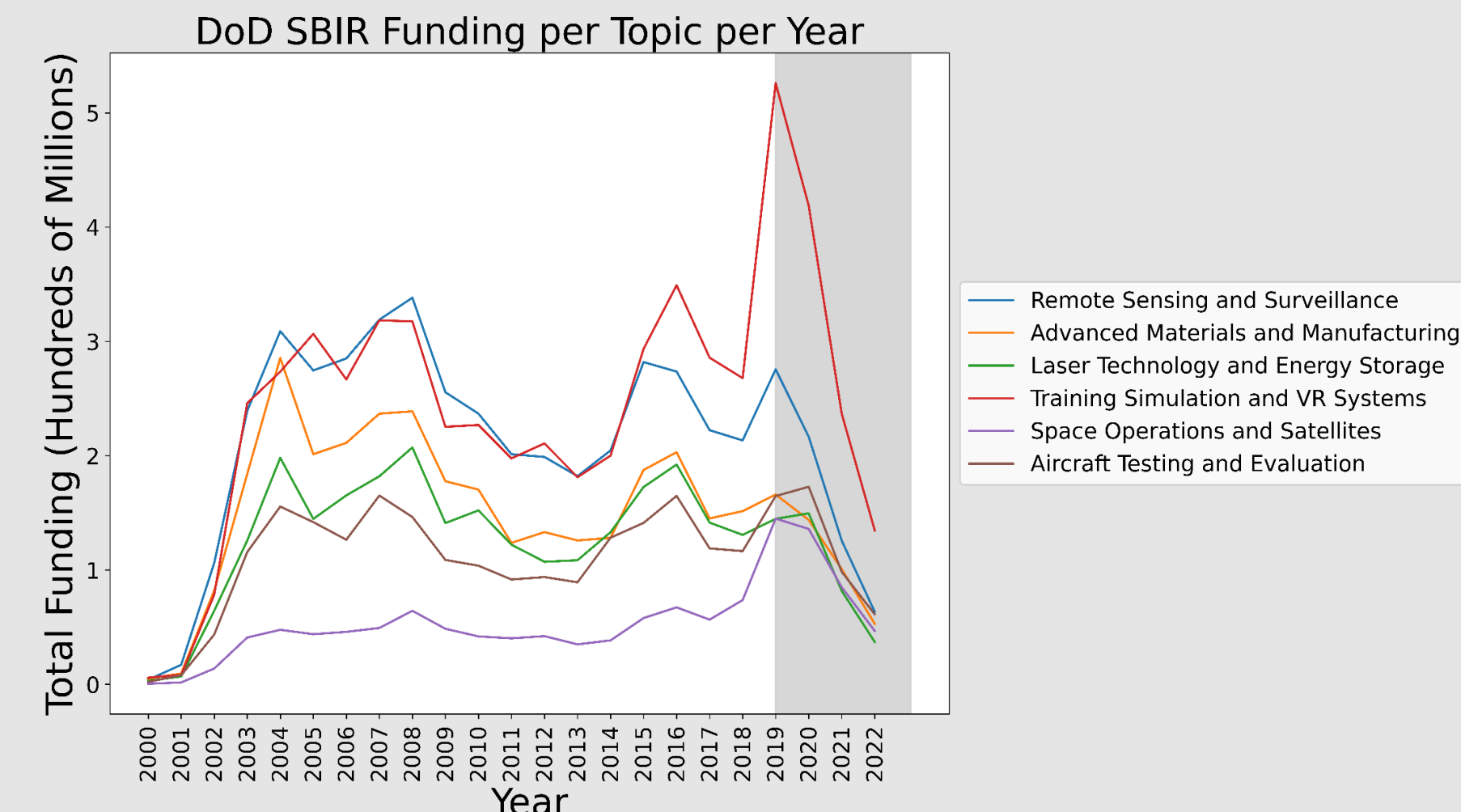
**Problem:** Emerging technologies are unprofitable until commercialization. It is very difficult to determine if a technology will transition or where the DoD is investing.

#### Our Approach:

- Split into two teams: Topic Modeling and Technology Readiness Levels (TRL) that analyze the Small Business Innovation Research (SBIR) program awards.
- Topic Modeling:** unsupervised method that converts award data into word clusters to seek patterns/trends.
- Technology Readiness Level Prediction:** Supervised method that uses BERT embeddings and upsampling to estimate technology maturity.

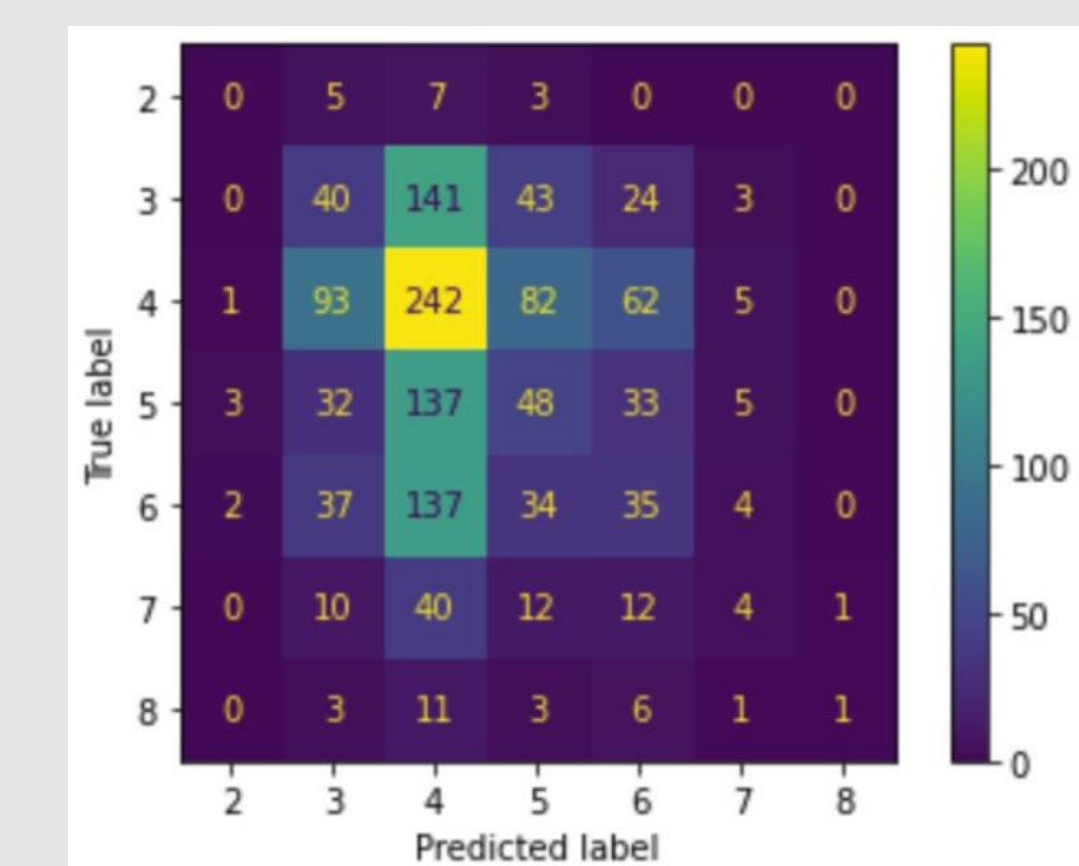
## RESULTS

### TEMPORAL TOPIC MODELING



Each line in the graph represents one type of topic that the DoD has allocated funding for. The total funding in hundreds of millions is represented by year (2000-2022).

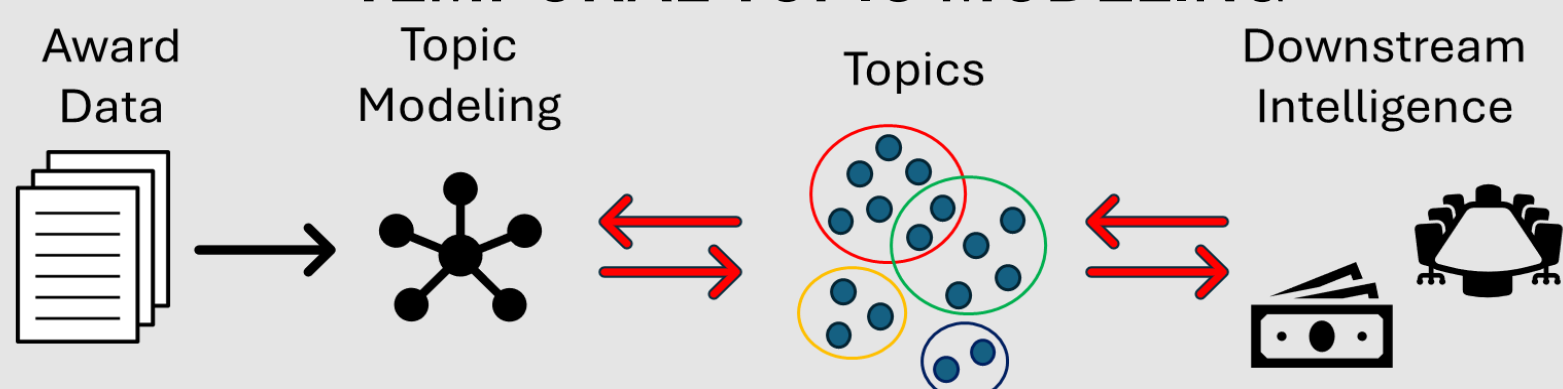
### TECHNOLOGY READINESS LEVEL PREDICTION



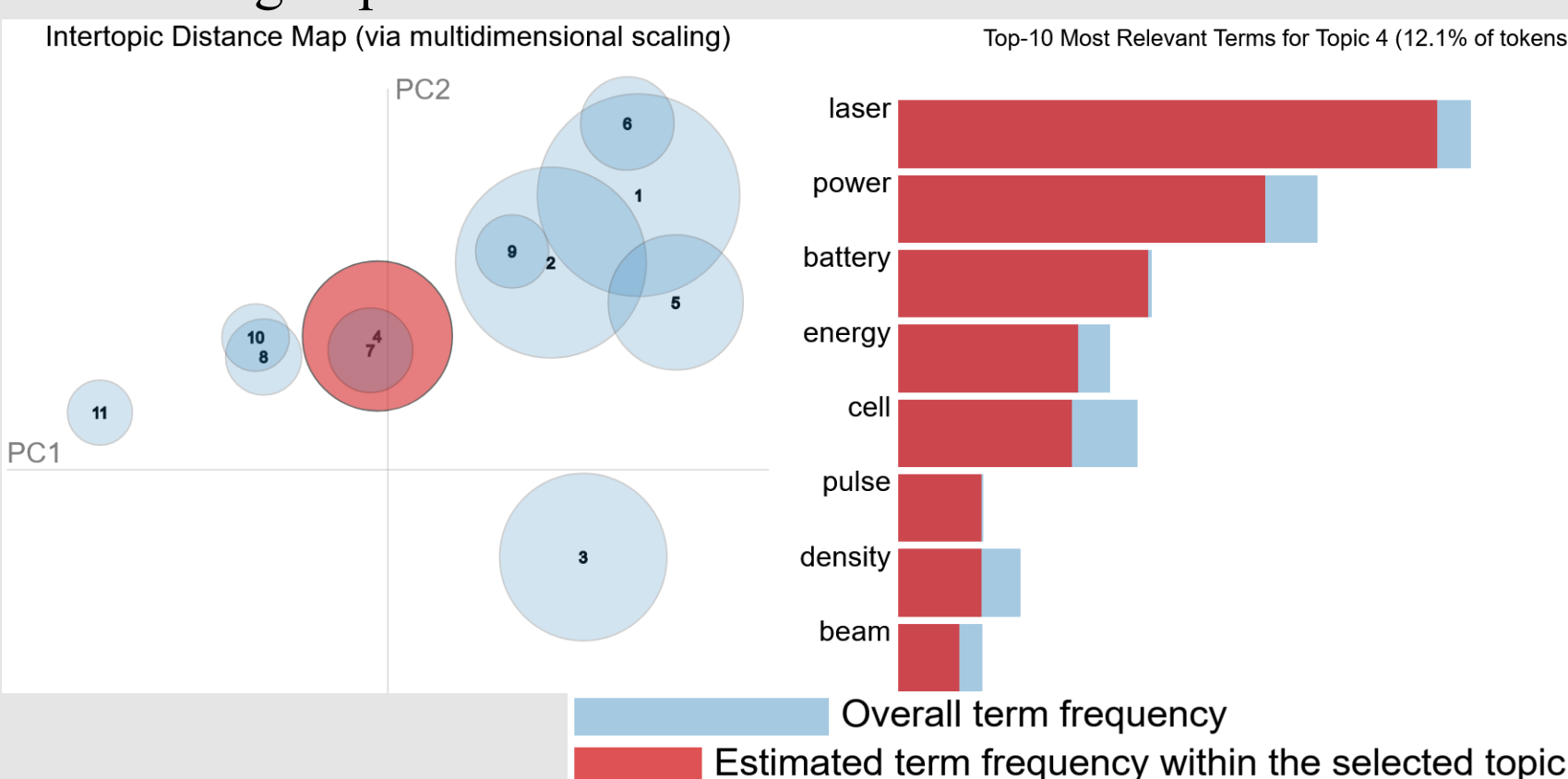
Confusion Matrix showing predicted label versus true label. The model is highly uncertain around TRL 4. This represents the transition from feasibility to building a prototype.

## METHODS

### TEMPORAL TOPIC MODELING

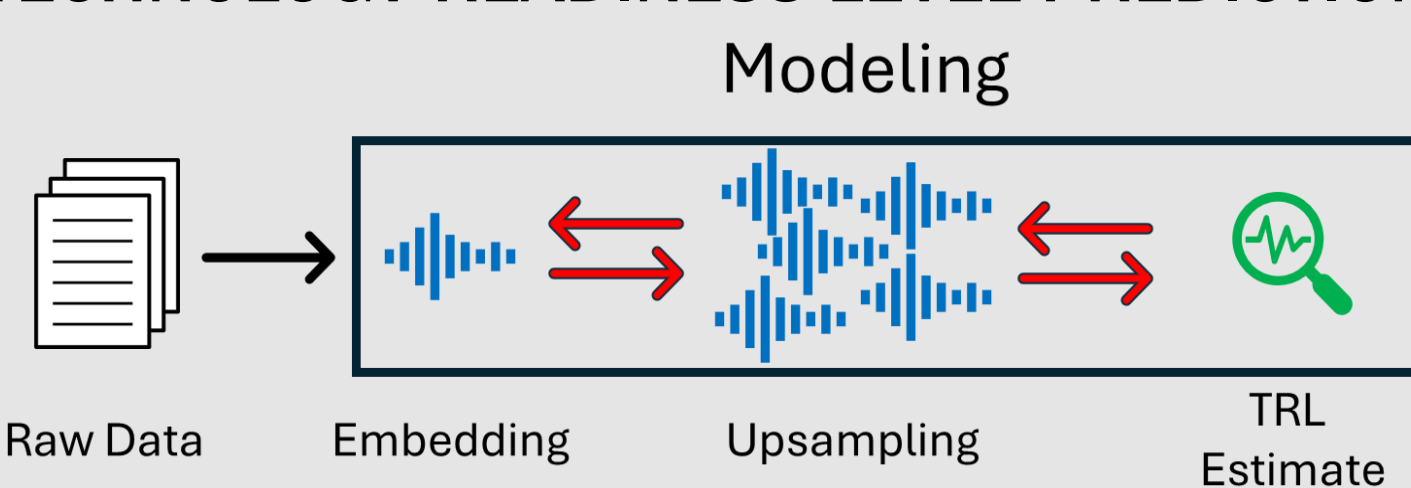


- Preprocessing** – remove noisy words and standardize text.
- Topic Modeling** – used to identify the main themes or topics within a set of text data.
- Topic Labeling** – refine parameters to get useful topics.
- Downstream Intelligence** – Topics are combined with funding to produce trends over time.

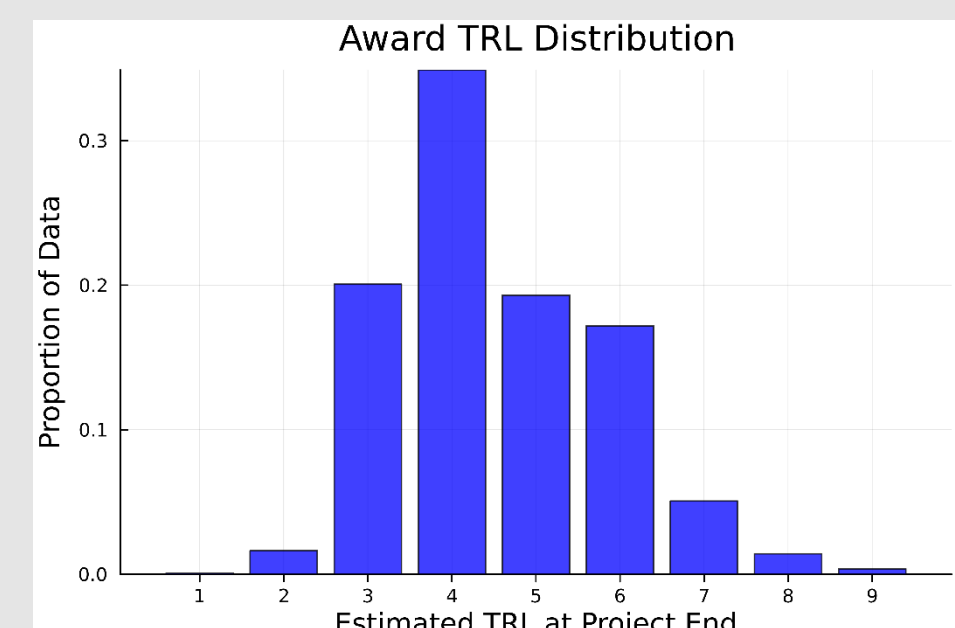


A static topic model showing topics and the word distribution for a single topic related to lasers and energy. Topic models can extract patterns in text but need dollars to realize business value.

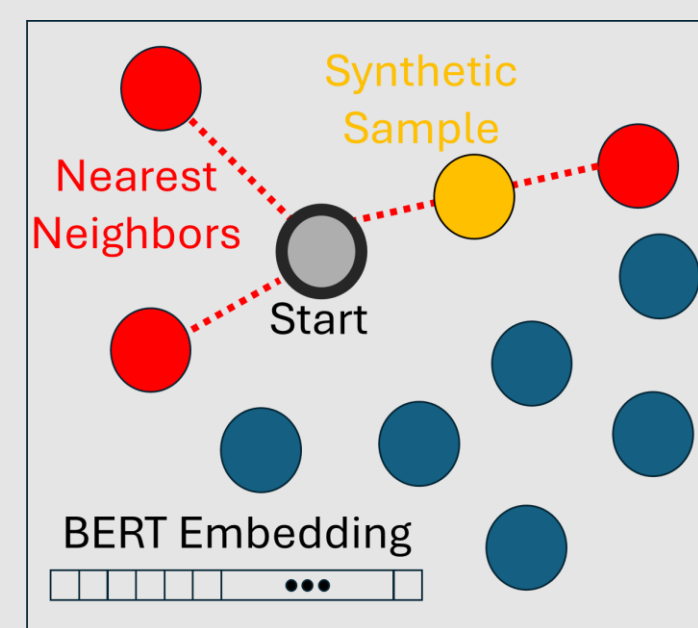
### TECHNOLOGY READINESS LEVEL PREDICTION



- Embedding** – BERT is used to preprocess and generate numerical representations of text data.
- Upsampling** – because we have limited data, we use SMOTE to upsample less frequent data.
- TRL Estimate** – To predict TRL, we experimented with various predictive models using upsampled embeddings.



Original Distribution. Higher TRL indicates a more mature product.



SMOTE is used to balance the data by generating synthetic samples from BERT embeddings.

## CONCLUSIONS

### Topic Modeling

Topics like space operations have seen increasing investments.

Other topics have experienced a decrease in SBIR funding (Advanced Materials and Manufacturing). However, this can be due to other funding avenues.

### TRL Prediction

Predicting TRL is very difficult problem. In some sense, we need to be able to store “the state of the world” implicitly in the embedding architecture.

The model could use more sophisticated embeddings and better strategies for handling the data imbalance.

## FUTURE WORK

### Topic Modeling

- More targeted topic analysis
- Improved topic interpretability
- Hierarchical topic models in each year
- Temporal granularity adjustment

### TRL

- LLMs – can capture prior knowledge
- Pretrain embeddings to fine-tune for our task
- Generative methods for data augmentation like back translation, or generative summaries.

## ACKNOWLEDGMENTS

We'd thank our mentor **Mike Douglass** for his time, support. We'd also like to thank other Raytheon staff that gave their time including **Karl Voigt, Colin Funai, Derrick Joyer, and Damianos Karakos**. We'd like to thank **Lauren Terese, Kali Lacy**, and the rest of The Data Mine staff for their support and encouragement. We thank the students from Fall 2023. Lastly, we'd like to acknowledge **Norma Grubb**. Part of our Topic Modeling code is based on examples she provided us in the 2023-2024 academic year.

## REFERENCES

Vally of Death in Defense Innovation Acquisition (Top leftmost graphic) Source: <https://media.defense.gov/2023/May/09/2003218877/-1/-1/0/NDSTS-FINAL-WEB-VERSION.PDF>