National Nuclear Security Administration

Management & Budget

Advanced Natural Language Processing for Work Breakdown Structures

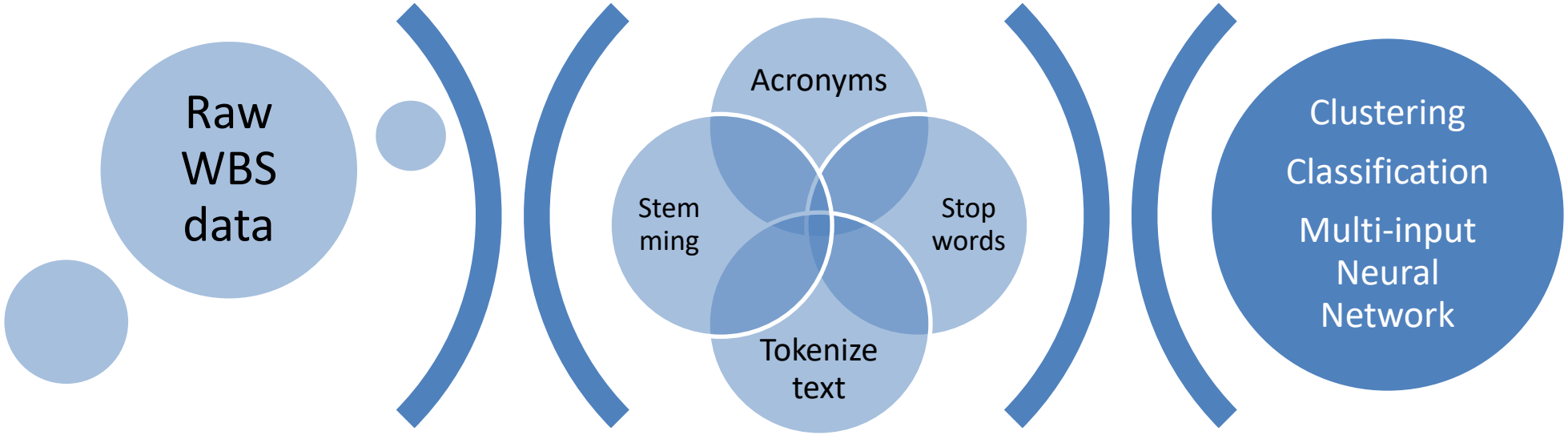
Maura Lapoff, Technomics supporting
NNSA Office of Programming, Analysis, and Evaluation (PA&E)

2022 ICEAA Professional Development and Training Workshop

The Problem and Solution

- The National Nuclear Security Administration (NNSA) collects Work Breakdown Structure (WBS) data for Capital Asset Projects.
- A WBS allows Cost Estimators and Program Managers to track and compare capital acquisition costs across the entire Nuclear Security Enterprise (NSE).
- Implementing the structure across projects grows exponentially in complexity due to varying project scope, contextual changes, and vendor requirements.
- Can computers assist in standardizing a Work Breakdown Structure?
  - Natural Language Processing (NLP) is a technical field that intersects linguistics, computer science, and machine learning (ML). The goal is to help computers process and analyze large amounts of data concerning humans' use of languages.
  - Useful for:
    - Classifying text (WBS elements, Documents, Your Tweets)
    - Information Extraction
    - Sentiment Analysis
    - Search Engines
    - Speech Recognition

Approach

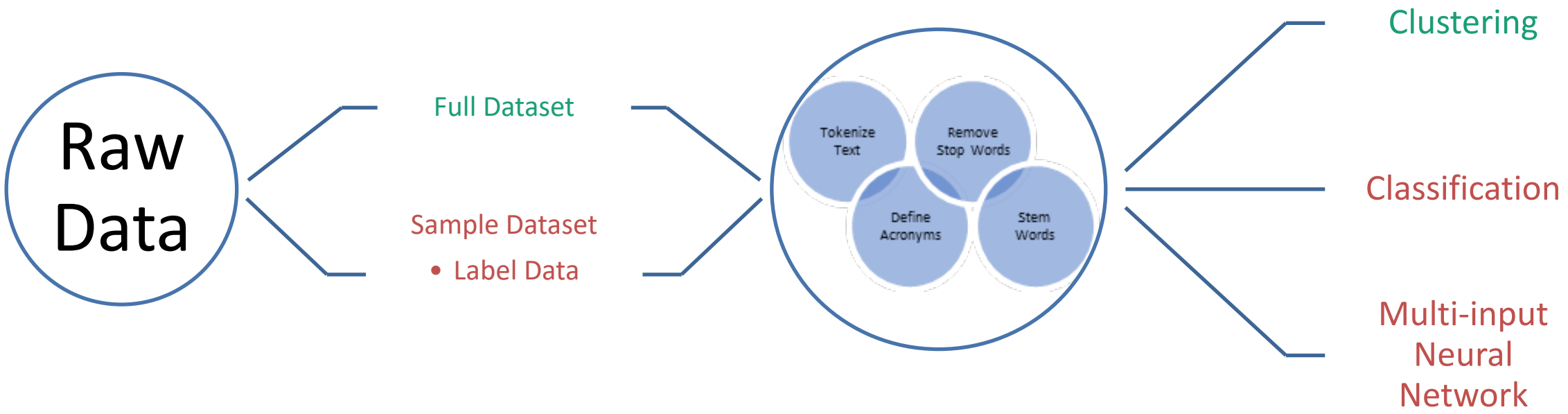


Export & Labelling

Preprocessing

Analysis

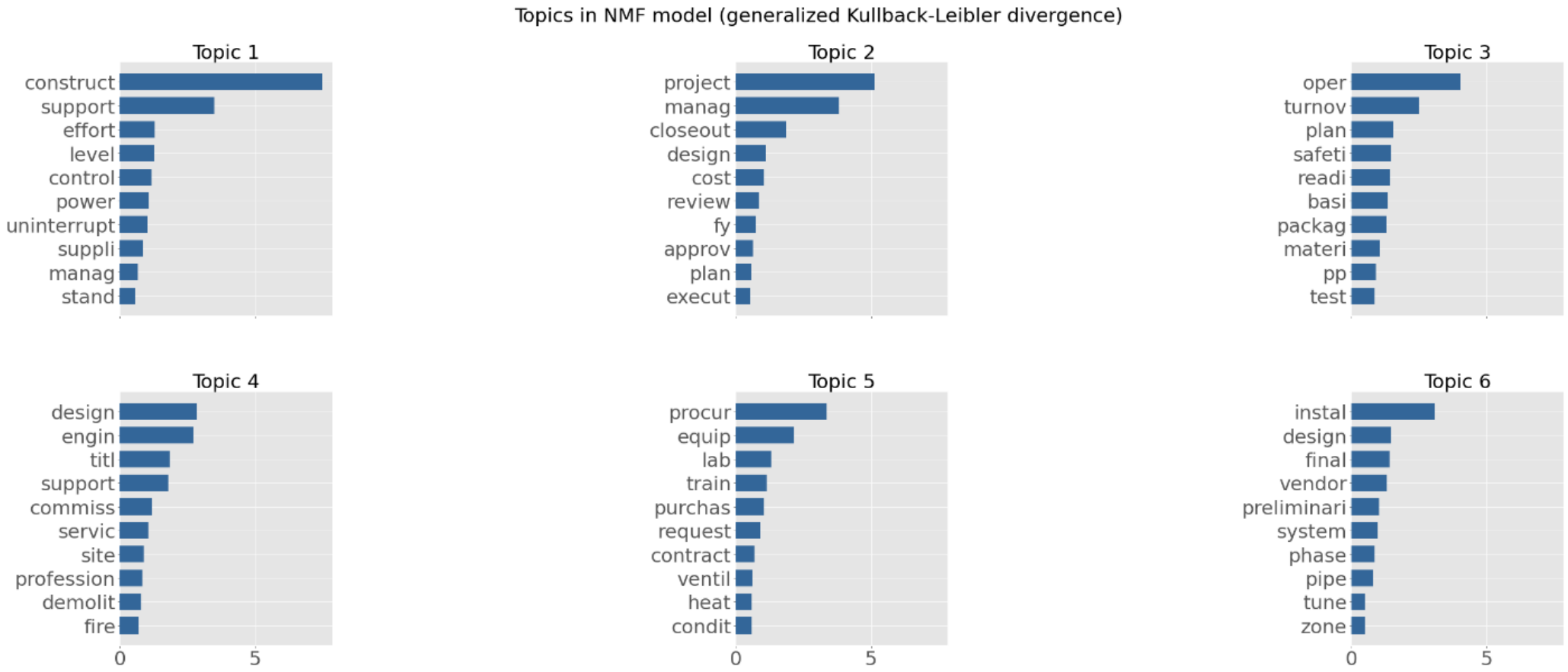
Data Preprocessing



Unsupervised Learning

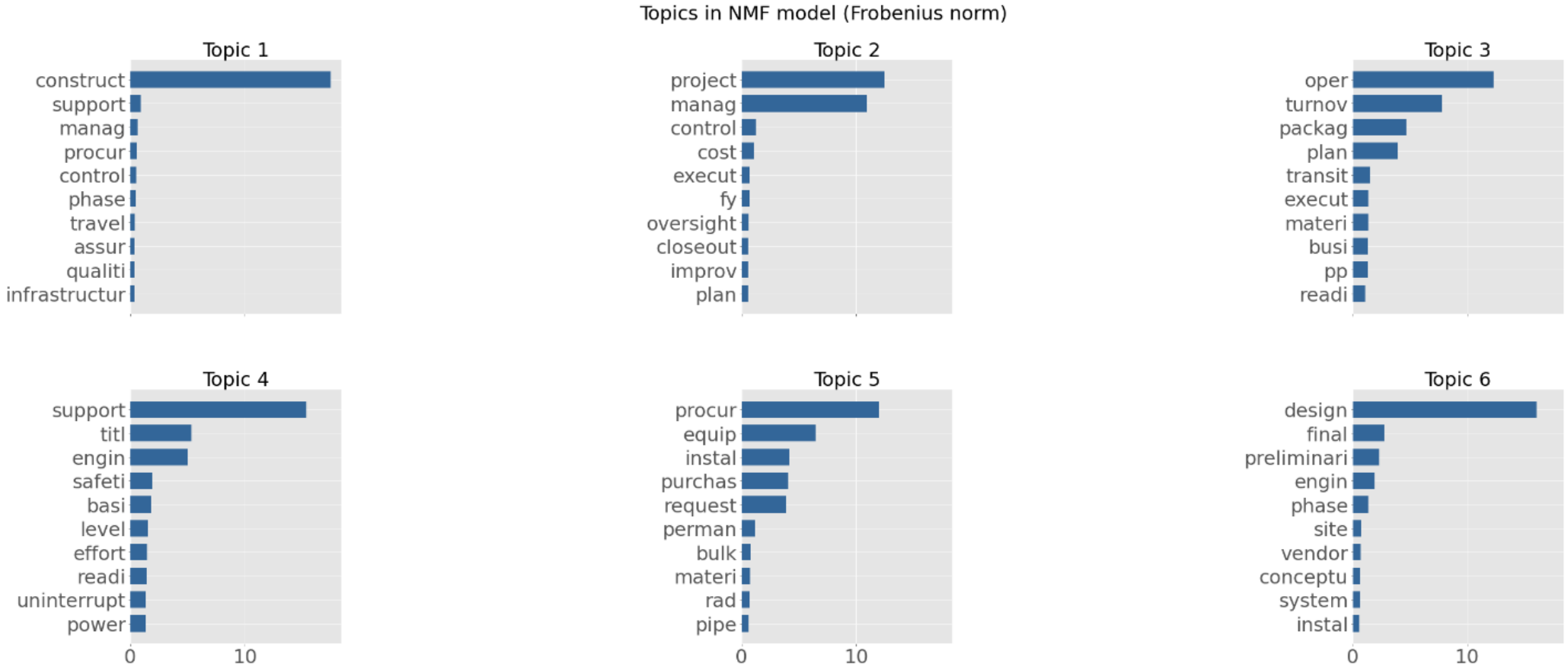
- Un-supervised – Clustering
  - Provide the full dataset and a set of Machine Learning Algorithms
    - Ask the computer to try and find natural clusters within the dataset
    - Select the number of clusters *k* to search for: **6**
  - No training – these are clusters that the computer thinks will be representative of the data that was used as input

Non-negative Matrix Factorization (NMF) with Kullback-Leibler divergence



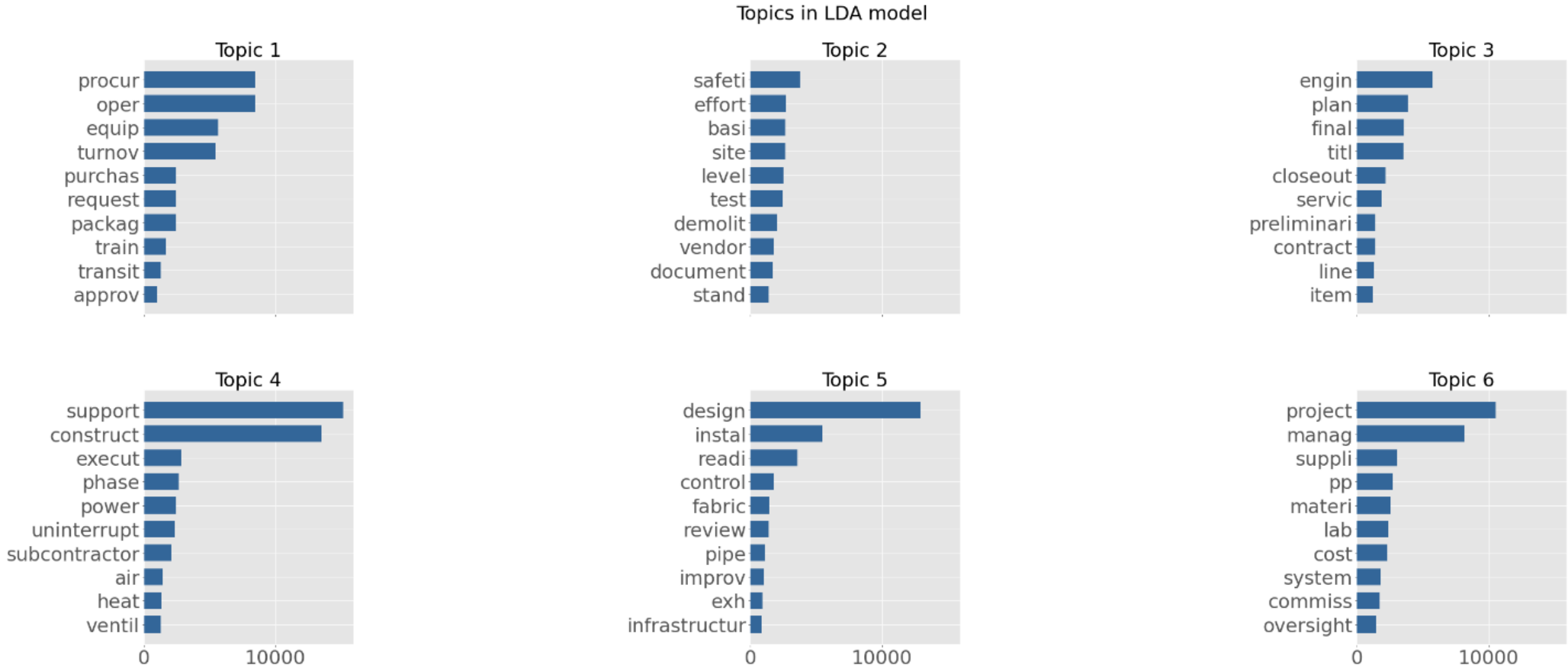
NMF-KL model validates the scheme: construction, project management, turnover to operations, design/engineering, procurement, and installation/design as the main topics

Non-negative Matrix Factorization (NMF) with Frobenius norm



Like the previous model, NMF-F model identifies construction, project management, turnover to operations, and procurement as topics; however, this model identifies support as its own topic and design is only associated with engineering and not installation

Latent Dirichlet Allocation

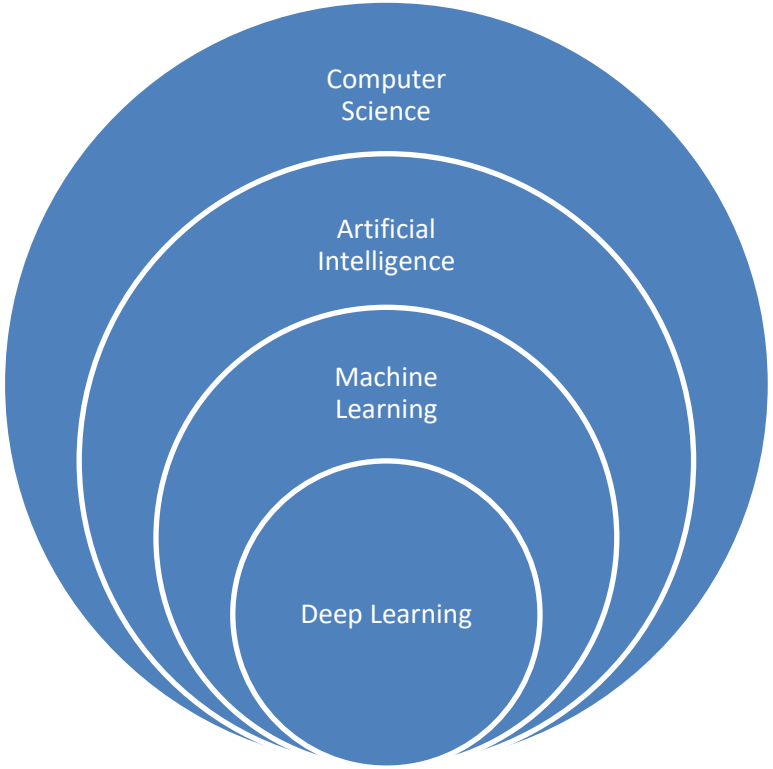


LDA model also shows procurement, project management, and design/installation as topics

Now, support is associated with construction, engineering is associated with planning, and there is a new cluster on safety basis/level of effort

Deep Learning

- Deep Learning is a sub-branch of Machine Learning dealing with the learning of representations and using those to make predictions.



Supervised Learning

- Supervised: Classification
  - Requires a training data set with our classification scheme.
    - Chose 6 Classes:
      - Construction, Project Engineering & Design, Procurement, Site Preparations, Start-Up, and Project Management
  - 80/20 training/testing split
  - Model Building:
    - Which algorithm produces the best results?
    - For a given algorithm, what set of input parameters produces the best results?
      - Model Tuning helps find the best values
    - Can we combine models to improve results? Which ones?

- Warning: Human error → error in results
  - Know your domain

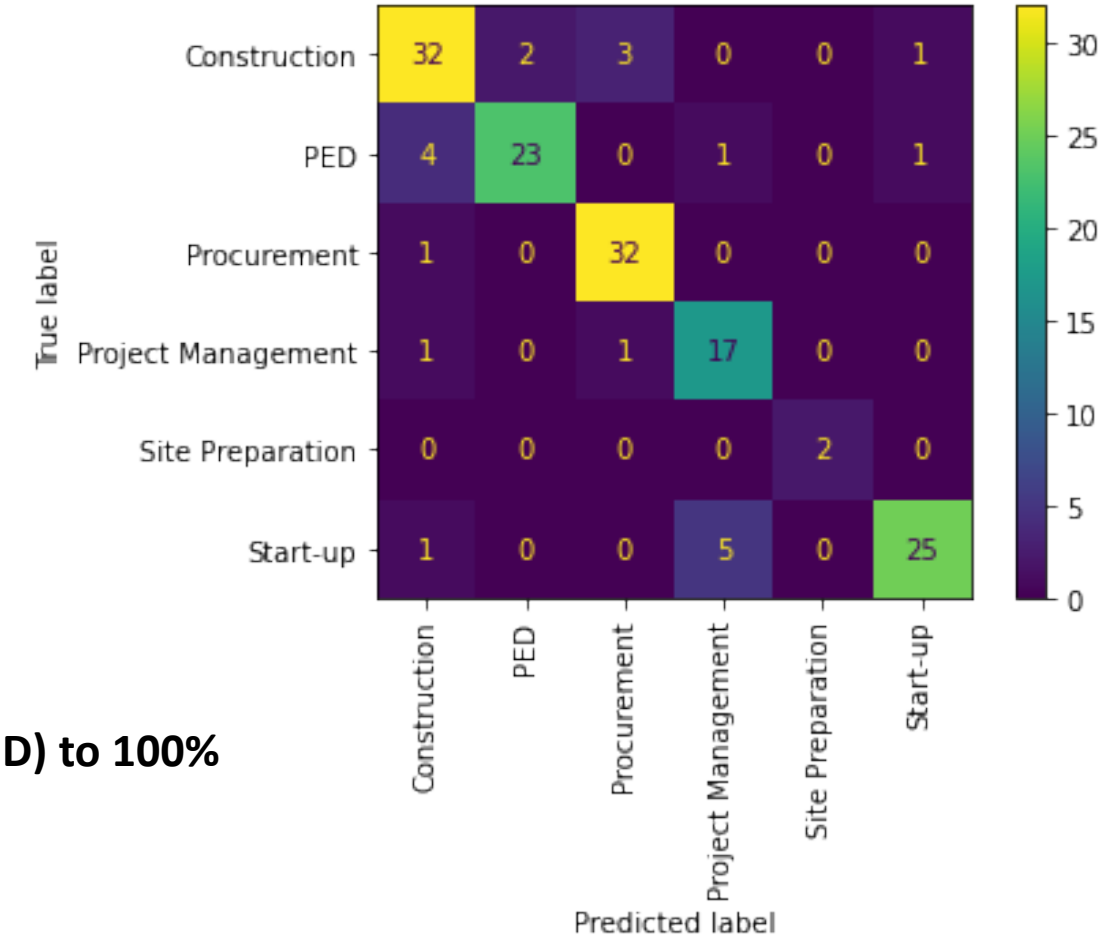
Tuned Model Results

- Three standard neural networks and two ensemble classifiers tested
- The best model was an ensemble model that weighted the Bi-LSTM and CNN models higher than the DNN

Classifier - Tuned	Accuracy
Densely Connected Neural Network	78%
Convolutional Neural Network	78%
Bidirectional LSTM	82%
Ensemble Voting – Equal weights	82%
Weighted Ensemble Voting	84%

Class Predictions of the Ensemble Networks

Class	Accuracy
Construction	88.9%
PED	79.3%
Procurement	97.0%
Project Management	89.5%
Site Preparation	100.0%
Start-up	80.6%



Accuracy at the class-level varies from 79.3% (PED) to 100% (Site Preparation)

Multi-Input Neural Networks

- Multi-input neural network uses both text and quantitative (ACWP) data
- Due to different model architecture, unable to use a Voting Classifier for the Multi-Input NN

Classifier	Accuracy
Densely Connected Neural Network	88.5%
Convolutional Neural Network	91.0%
Bidirectional LSTM	93.1%
Ensemble Voting	NA

- Using both types of inputs means that best model performance increases from 84% to 93%.

Observations and Conclusion

- Results of clustering models generally validate the use of the 6 classification labels used in the classification models
  - The computer and the analyst struggled with similar concepts, i.e., whether design and engineering belong together and whether support should be its own category
- Limitations
  - Acronyms
  - Stop Words
  - Classification Scheme
  - Sample Selection
  - Availability of Data, especially Cost Data

This advance means we can better track and compare project costs within and between projects, improving understanding of these projects and enabling more informed project and portfolio management decision-making for capital asset projects.

References

1. Géron, Aurélien. 2017. *Hands-on machine learning with Scikit-Learn and TensorFlow: Concepts, tools, and techniques to build intelligent systems.* O'Reilly Media.
2. Department of Energy Office of Project Management. 2021. "DOE Order 413.3B: Program and Project Management for the Acquisition of Capital Assets." *DOE Directives.* Washington, DC: Department of Energy, January 12. https://www.directives.doe.gov/directives-documents/400-series/0413.3-BOrder-b-chg6-ltdchg/@@images/file.

Backup

Multi-input Bi-LSTM Architecture