INNOVATE. COLLABORATE. DELIVER.

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

National Nuclear Security Administration

Presented at the 2022 ICEAA Professional Development & Training Workshop: www.iceaaonline.com/pit2022

Bottom Line Up Front

- The National Nuclear Security Administration (NNSA) collects Work Breakdown Structure (WBS)-format data for capital asset projects.
- Capital Assets are defined in DOE Order 413.3B, Program and Project Management for the Acquisition of Capital Assets.
- We applied advanced unsupervised and supervised machine learning methods to validate and classify NNSA Capital Asset projects to a common Level 2 WBS.

NNSA Mission

Protect the Nation by maintaining **a safe, secure, and effective** nuclear weapons stockpile

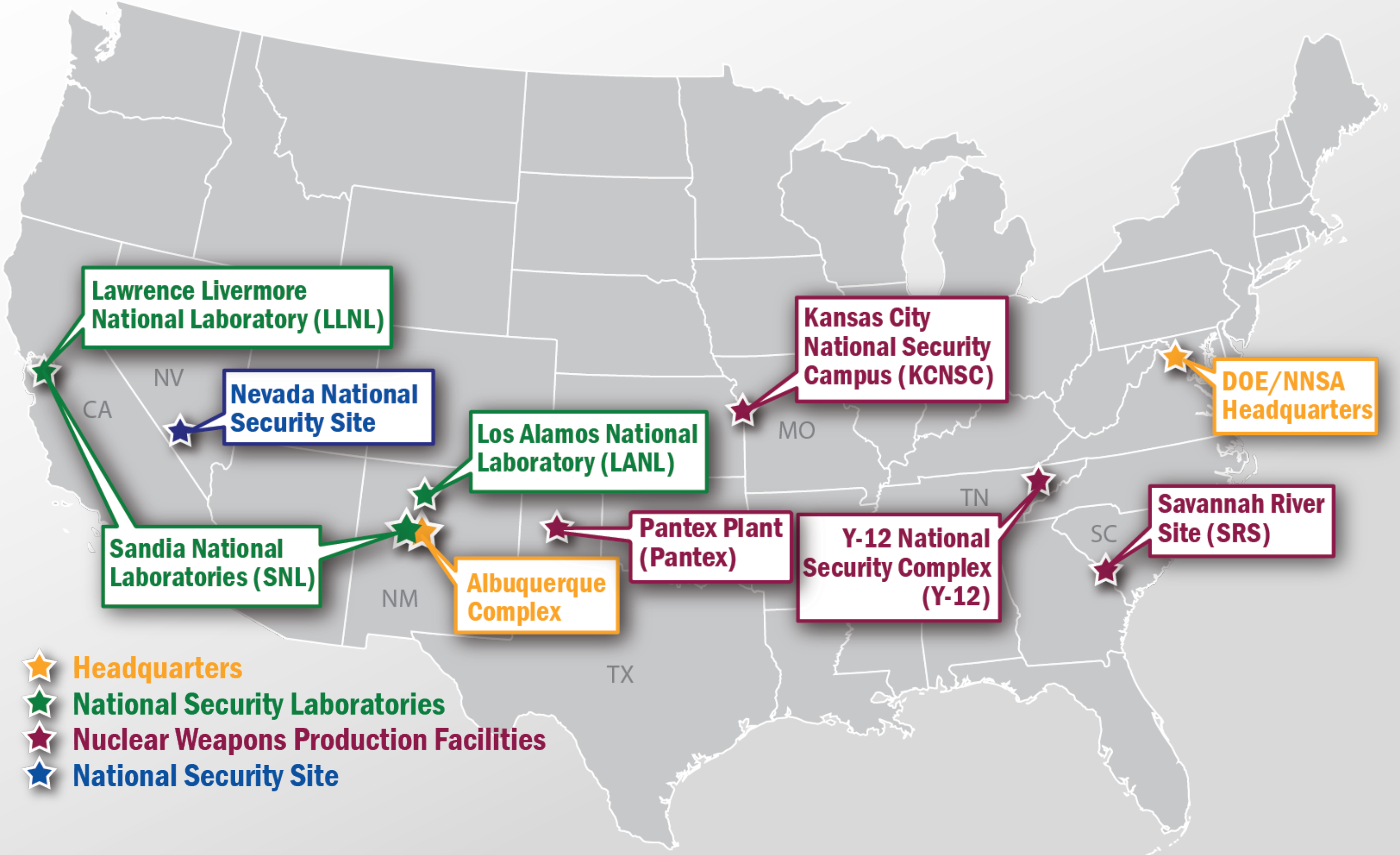
Reduce global **nuclear threats**

Provides the U.S. Navy with militarily effective **nuclear propulsion**

The NNSA is an agency within the U.S. Department of Energy responsible for enhancing national security through the military application of nuclear science

NNSA Labs, Plants, and Sites

INNOVATE. COLLABORATE. DELIVER.



NNSA Infrastructure

- Most facilities are **over 40 years old**, nearly 30% date to the Manhattan project era, and nearly two-thirds are in less than adequate condition
  - This age and condition creates a challenge to safely operate and meet mission demands
- Projects include:
  - Cutting-edge research laboratories
  - Safe and secure weapon assembly plants
  - Nuclear and non-nuclear component manufacturing
  - Office facilities

NNSA Capital Projects range from office buildings to the National Ignition Facility (Pictured Above)

The Problem and Solution

- The National Nuclear Security Administration (NNSA) collects cost data organized according to a Work Breakdown Structure (WBS) 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

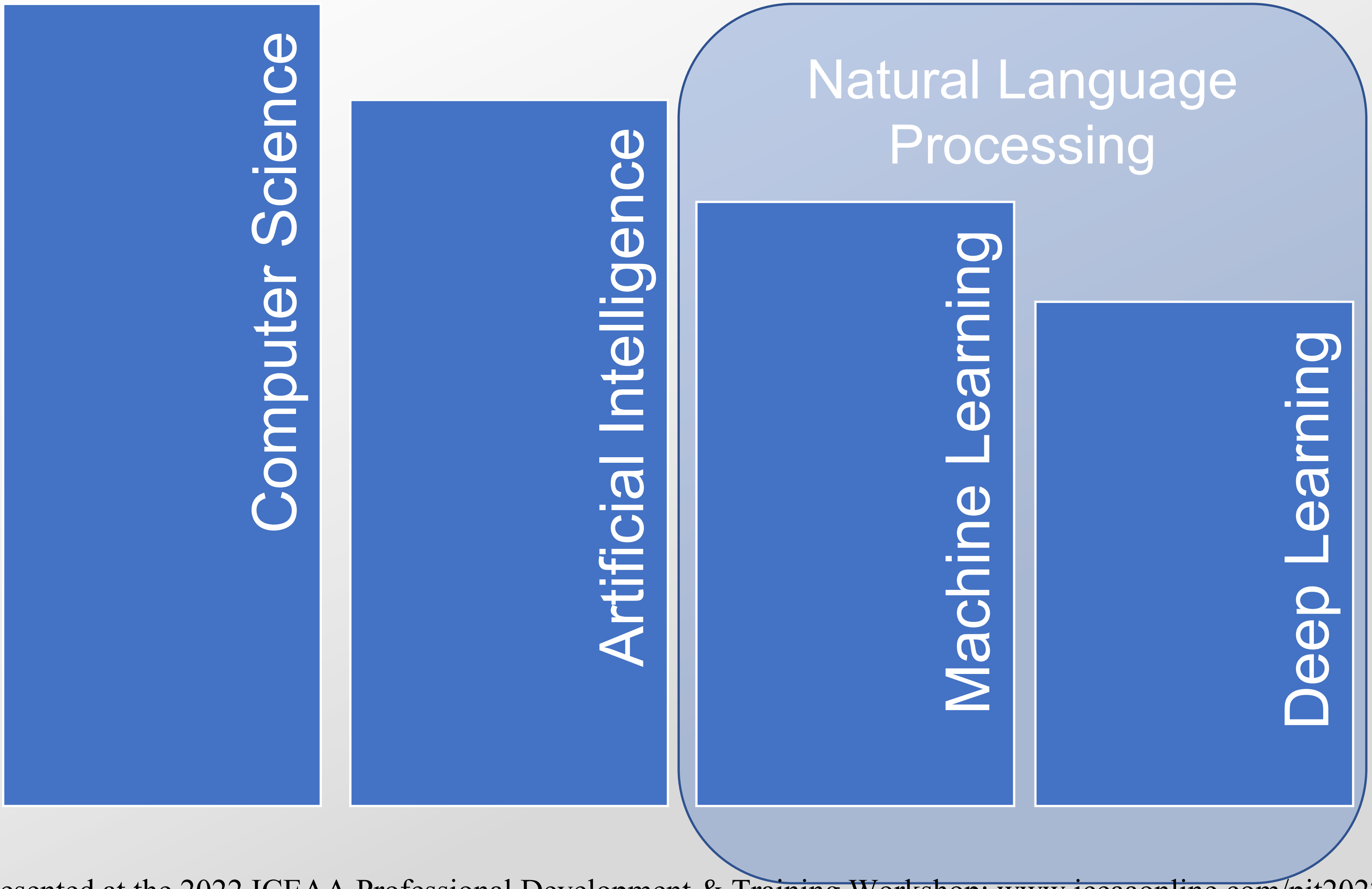
Terminology

Machine Learning	Natural Language Processing
Deep Learning	Artificial Intelligence

?

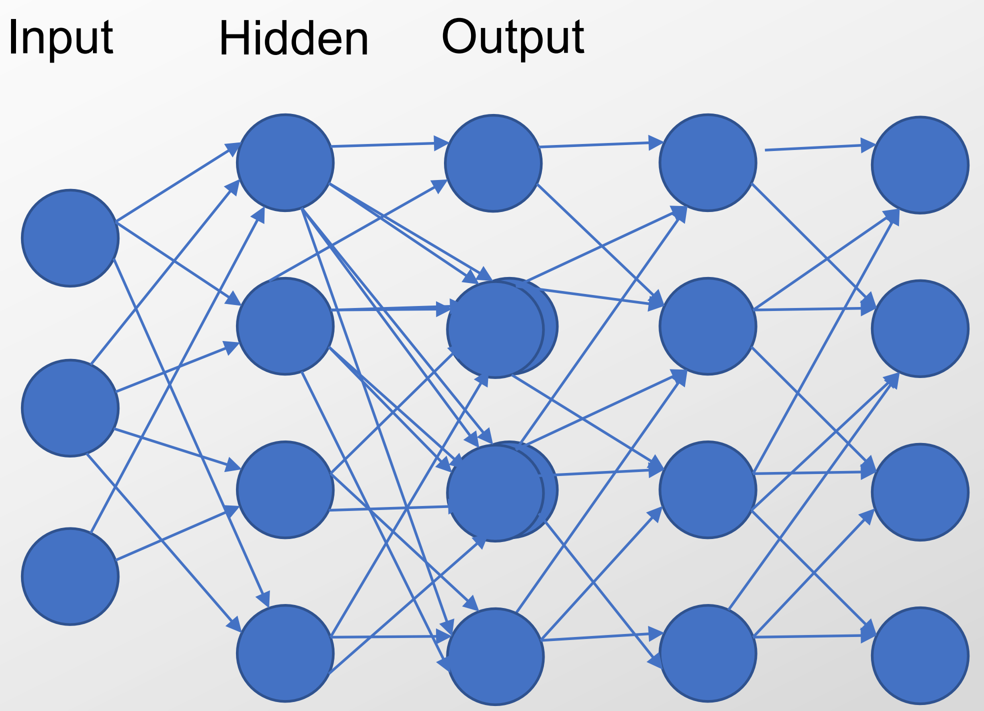
Discipline Hierarchy

INNOVATE. COLLABORATE. DELIVER.



Artificial Neural Networks

INNOVATE. COLLABORATE. DELIVER.



Multi-Layer Perceptron

- A basic MLP node *i* fires when $y_i > 0$.

$$y_i = \sum_{i=1}^{k} w_j x_j + b \qquad \text{(a)}$$

$$f(x) = \left\{ \begin{array}{c|c} 1 & y_i > 0 \\ 0 & y_i < 0 \end{array} \right\} \qquad \text{(b)}$$

Approach

INNOVATE. COLLABORATE. DELIVER.

Raw WBS data

Export & Labelling

Acronyms

Stem ming

Stop words

Tokenize text

Preprocessing

Clustering

Classification

Multi-input Neural Network

Modeling & Analysis

Data Collection

INNOVATE. COLLABORATE. DELIVER.

- 147 Projects → 31 Projects
  - 9 Active
  - 18 Completed / Completed & Closed
  - 2 Cancelled
  - 2 Other

Data Preprocessing

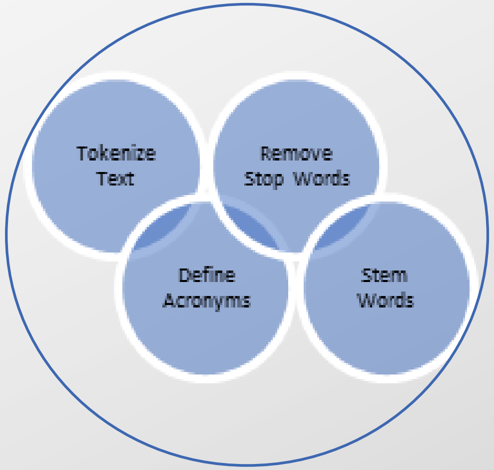
INNOVATE. COLLABORATE. DELIVER.

Raw Data

Full Dataset
(31 Projects)

Sample Dataset
(10 Projects)
- Label Data

1. Construction
2. Project Engineering and Design
3. Procurement
4. Project Management
5. Site Preparation
6. Start-up



Clustering

Classification

Multi-input Neural Network

Tokenization

- Tokenize Text

1	2	3	4	5	6	7	8	9	10	11
[The]	[cat]	[sat]	[on]	[the]	[ledge]	[and]	[stared]	[out]	[the]	[window]

Define Acronyms

- Define Acronyms
  - NYC, Manhattan, New York City → New York City
  - Washington DC, DC, District of Columbia, Washington, D.C. → Washington, DC
  - Do acronyms mean the same thing across (and within) project data?

Remove Stop Words

- Exclude words that are irrelevant or may bias results
  - Generic words: a, the, it, by, etc.
  - Domain-specific words: project names, company names, etc.

Stem Words

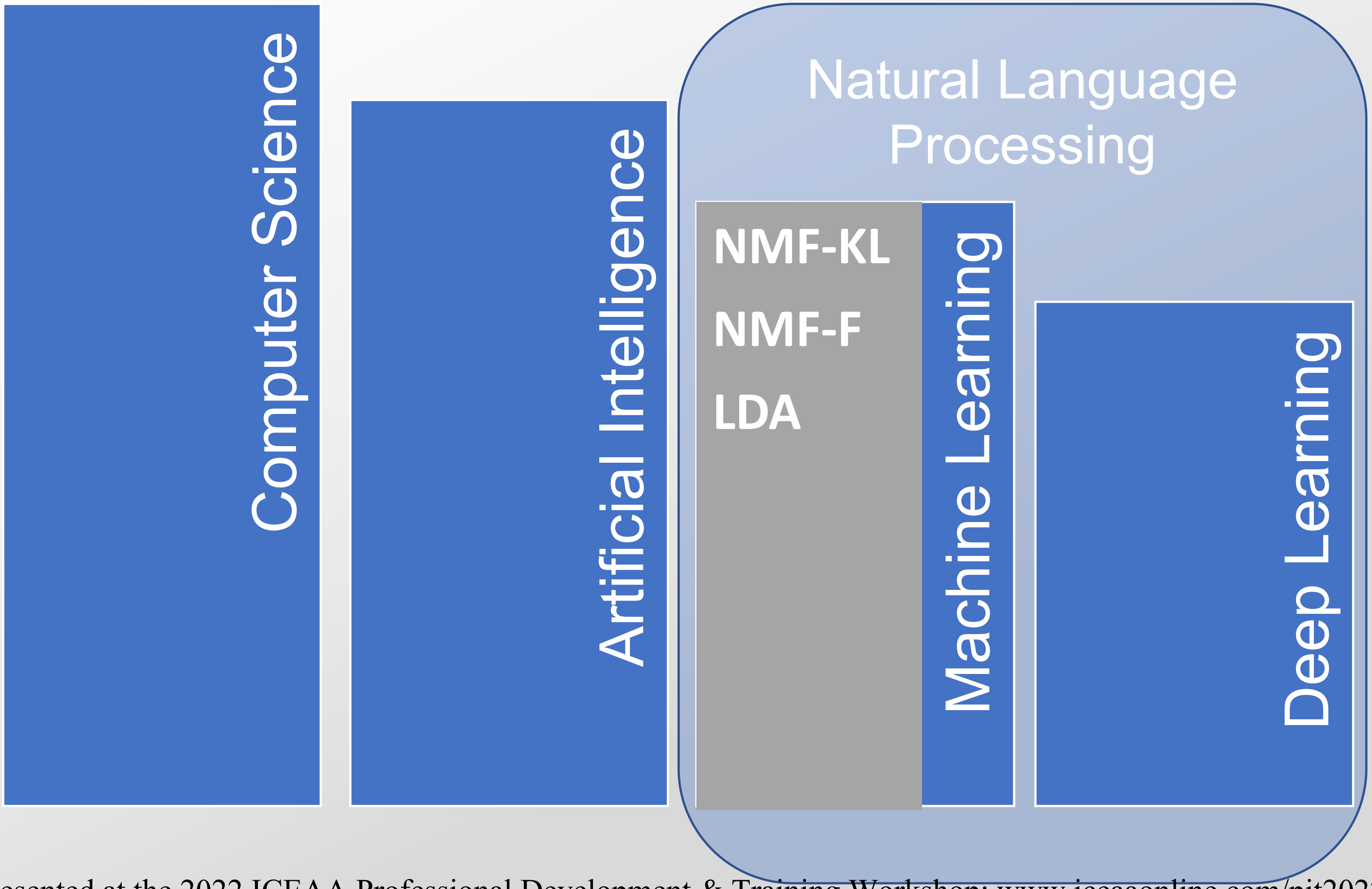
- (Optionally) reduce words to their common stem
  - Manage, Management, Managing → manag

Unsupervised Learning

- Un-supervised – Clustering
  - Provide the full, cleaned 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
  - Three algorithms:
    - Non-negative matrix factorization with Frobenius norm (NMF-F)
    - Non-negative matrix factorization with Kullback-Leibler Divergence (NMF-KL)
    - Latent Dirichlet Allocation (LDA)

Unsupervised Learning

INNOVATE. COLLABORATE. DELIVER.



Term frequency – inverse document frequency (tf-idf)

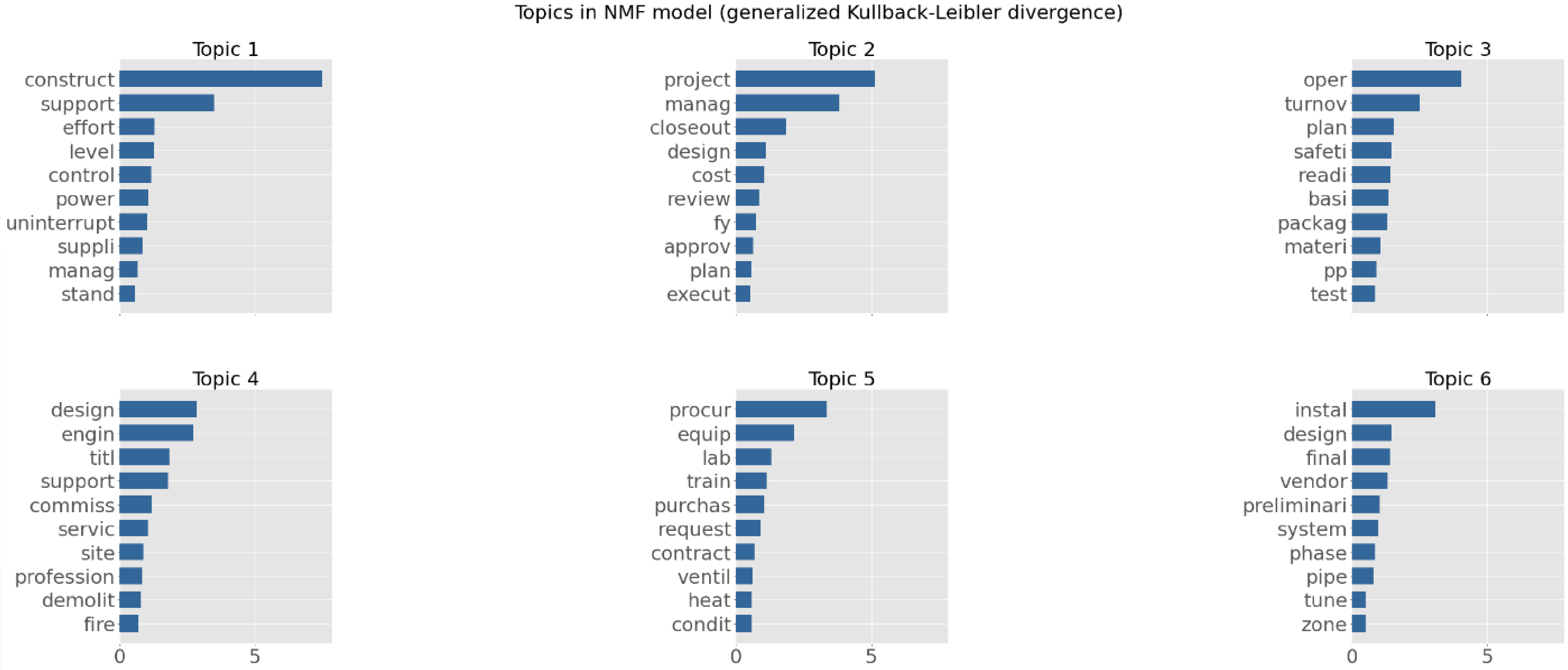
$$tf_{w,d} = \frac{\mathrm{n}_{\mathrm{w,\,d}}}{\sum_{\mathrm{k}} \mathrm{n}_{\mathrm{w,\,d}}}$$

$$idf_{w,D} = \log \frac{D}{df_w}$$

$$tf\text{-}idf = tf_{w,d} \cdot idf_{w,D}$$

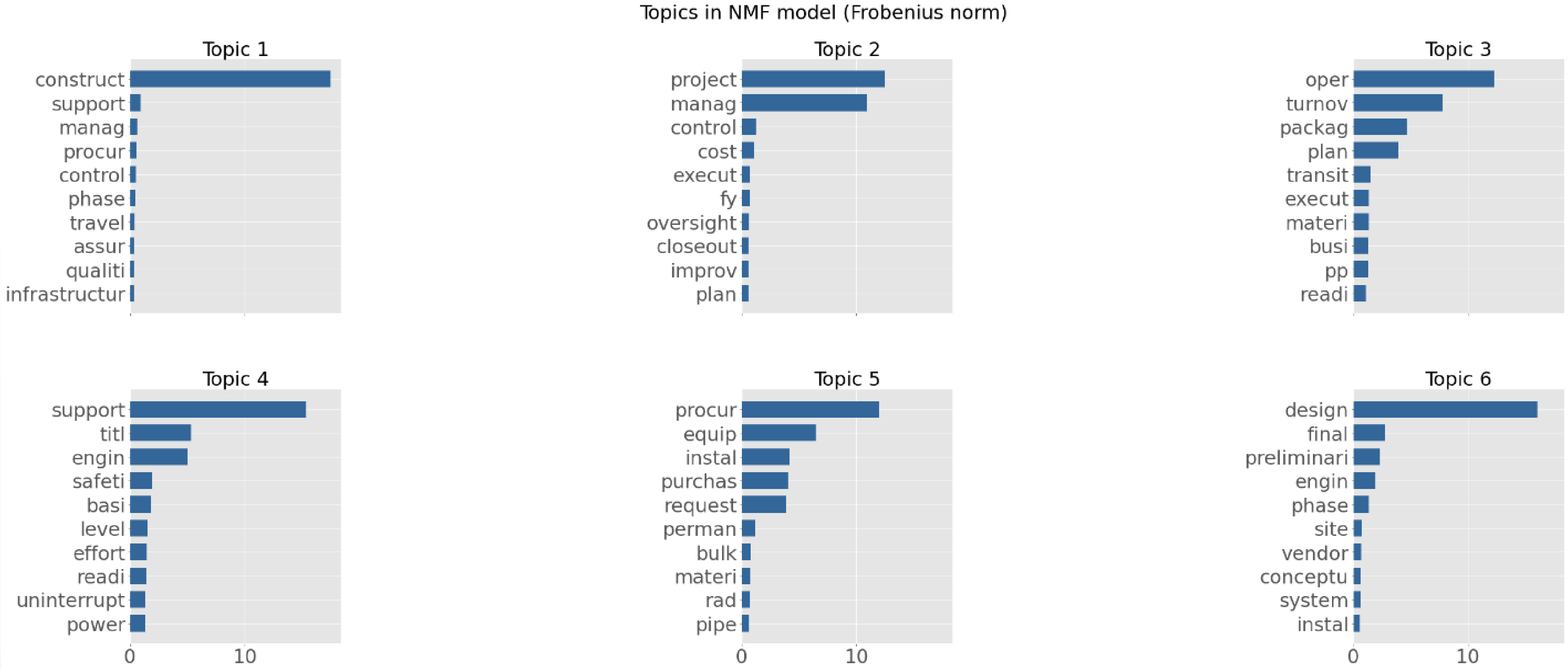
Non-negative Matrix Factorization (NMF) with Kullback-Leibler divergence

INNOVATE. COLLABORATE. DELIVER.



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

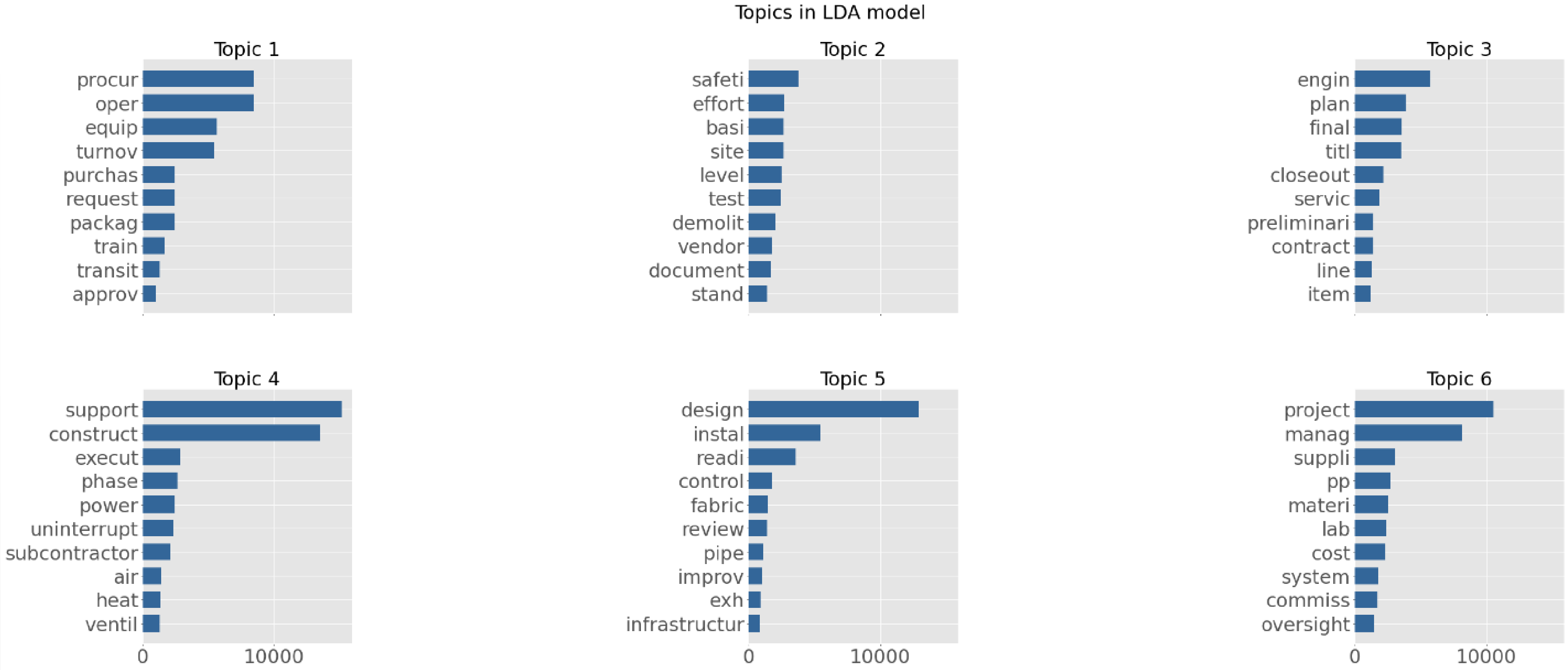
Non-negative Matrix Factorization (NMF) with Frobenius norm



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

INNOVATE. COLLABORATE. DELIVER.



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

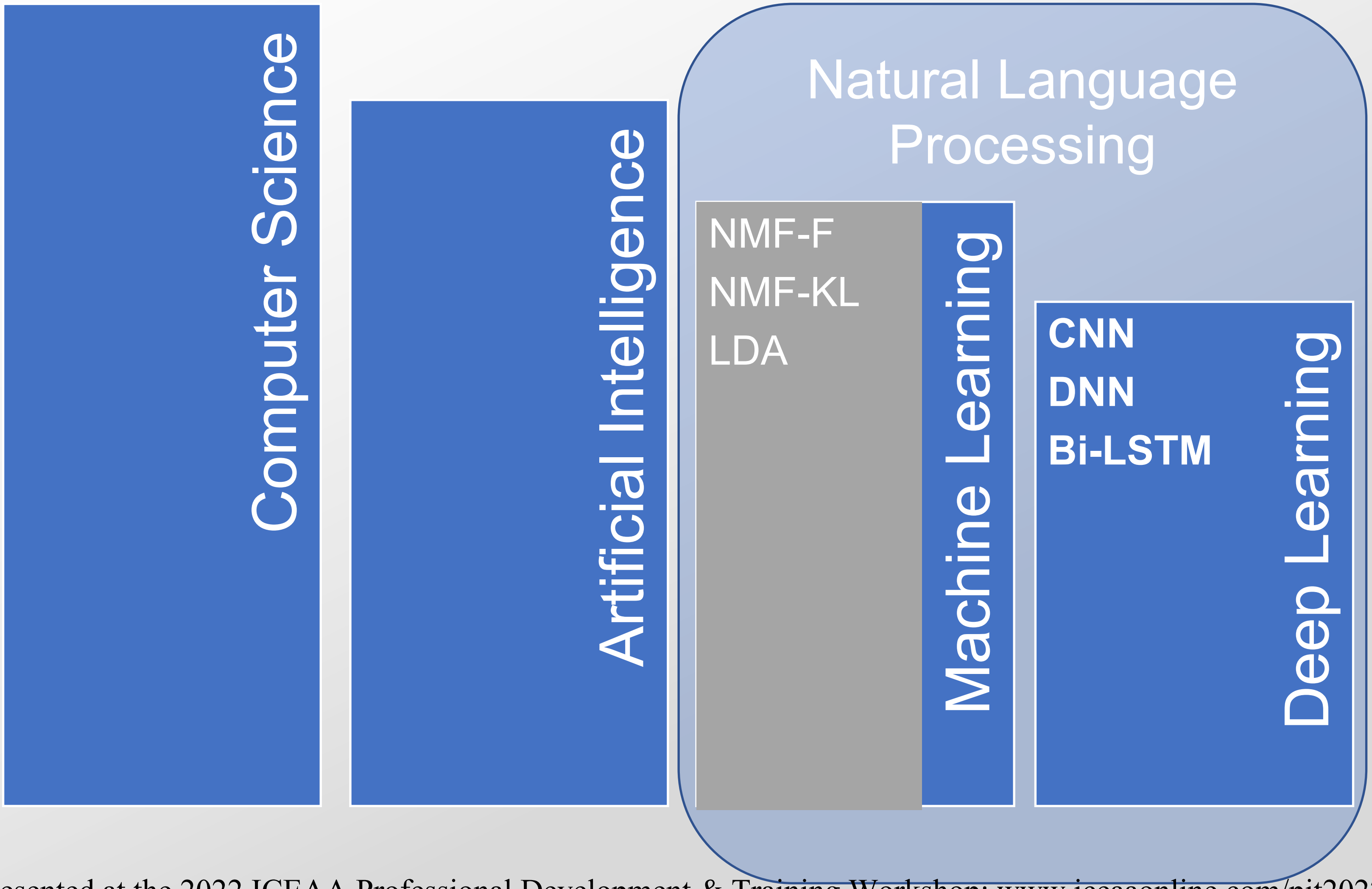
Supervised Learning

- 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

Supervised Learning

INNOVATE. COLLABORATE. DELIVER.



Model Tuning

- 2 parameters x 2 values each using three-fold cross-validation
  - 4 trials
  - 12 predictions

- 4 parameters x 5 values each with three-fold cross-validation
  - 1024 trials
  - 3072 predictions

Model Results

INNOVATE. COLLABORATE. DELIVER.

Classifier	Initial - Accuracy
Dense Neural Network	73.0%
Convolutional Neural Network	57.2%
Bidirectional LSTM	69.7%

Model Results

INNOVATE. COLLABORATE. DELIVER.

Classifier	Initial - Accuracy	Tuned - Accuracy
Dense Neural Network	73.0%	74.2%
Convolutional Neural Network	57.2%	72.6%
Bidirectional LSTM	69.7%	74.9%

Model Results

INNOVATE. COLLABORATE. DELIVER.

Classifier	Initial - Accuracy	Tuned - Accuracy	Retrained
Dense Neural Network	73.0%	74.2%	77.6%
Convolutional Neural Network	57.2%	72.6%	78.3%
Bidirectional LSTM	69.7%	74.9%	81.6%

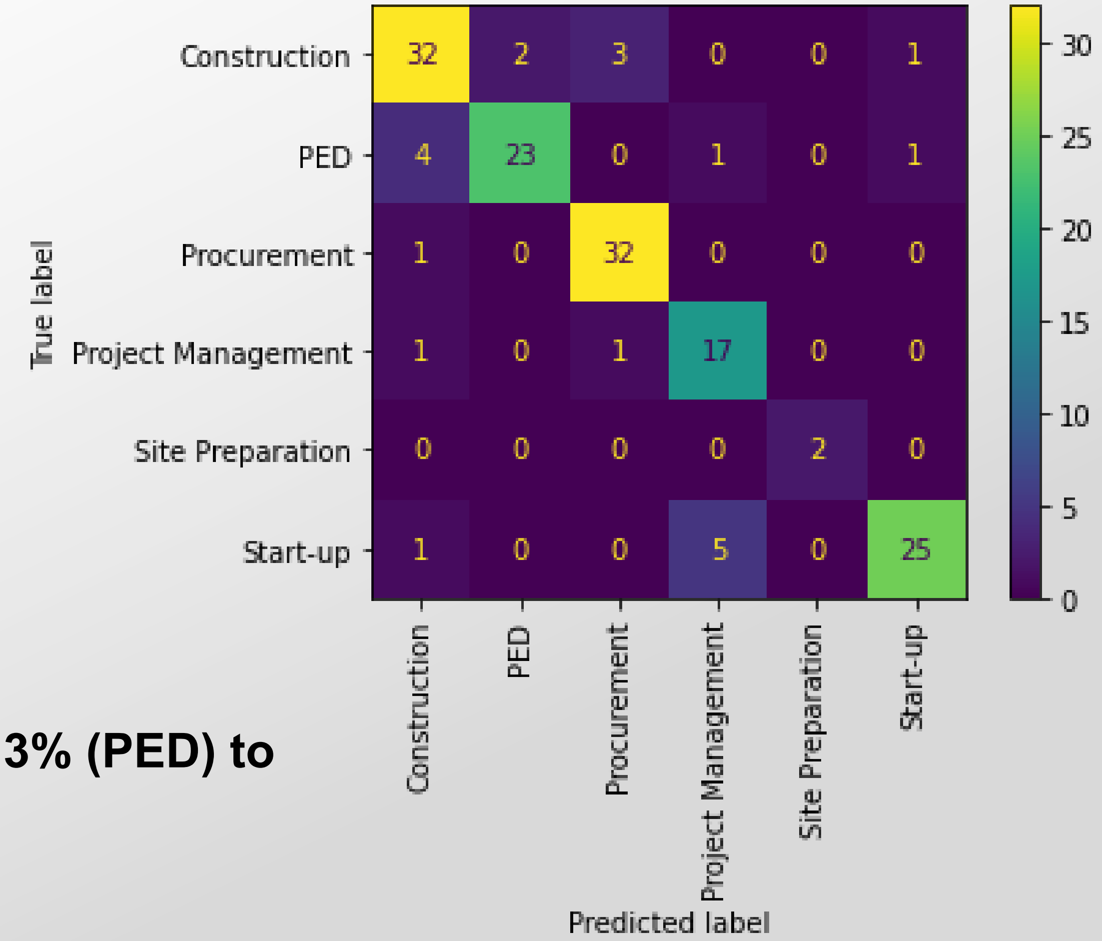
Model Results

Classifier	Tuned - Accuracy
Dense Neural Network	77.6%
Convolutional Neural Network	78.3%
Bidirectional LSTM	81.6%
Ensemble Voting – Equal weights	82.2%
Weighted Ensemble Voting	83.6%

- 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

Class Predictions of the Ensemble Network

Class	Accuracy
Construction	84.2%
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

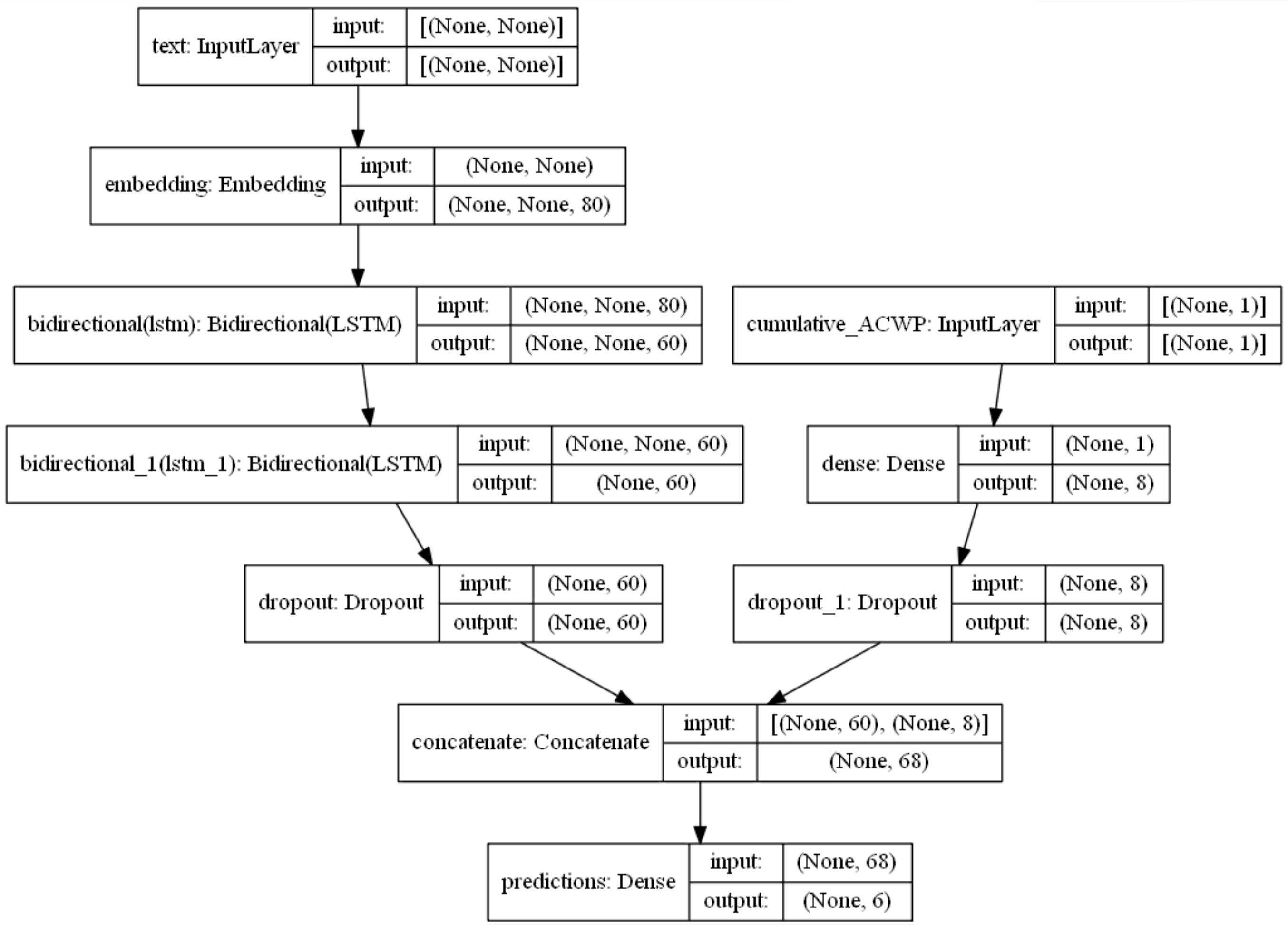
- Multi-input neural network uses both text and quantitative (ACWP) data
- Due to different model architecture, unable to use a Voting Classifier for the Multi-Input NN

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%.

Multi-input Bi-LSTM Architecture

INNOVATE. COLLABORATE. DELIVER.



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.

The full list of references can be found in the accompanying paper for this presentation.

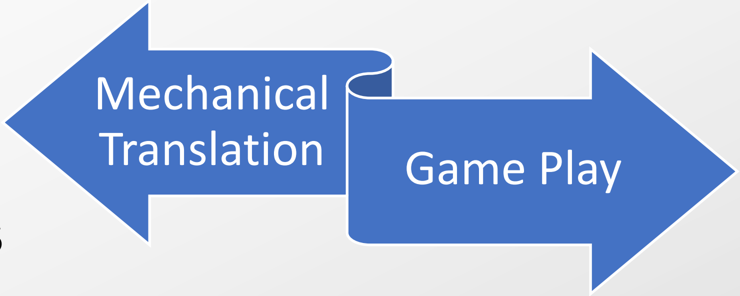
Thank you!

mlapoff@technomics.net

Backup

History

- 1952 – First Mechanical Translation (MT) Conference
  - Discussed complexities
  - No demos of true MT
- 1954 – First successful demonstration of Machine Translation



- 1952 – IBM checker program
- 1955 – Chess program can "learn" to improve performance