All in the Hierarchy: Meta-Estimators for Work Breakdown Structures

Maura Lapoff, Technomics

Patrick Shelton, Technomics

In Support of NNSA Office of Programming, Analysis, and Evaluation (PA&E)

2023 ICEAA Professional Development and Training Workshop

Agenda

- Introduction
- Overview
- Previous Work
- Current Work
- Conclusion
- References

NNSA Mission

INNOVATE. COLLABORATE. DELIVER.

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

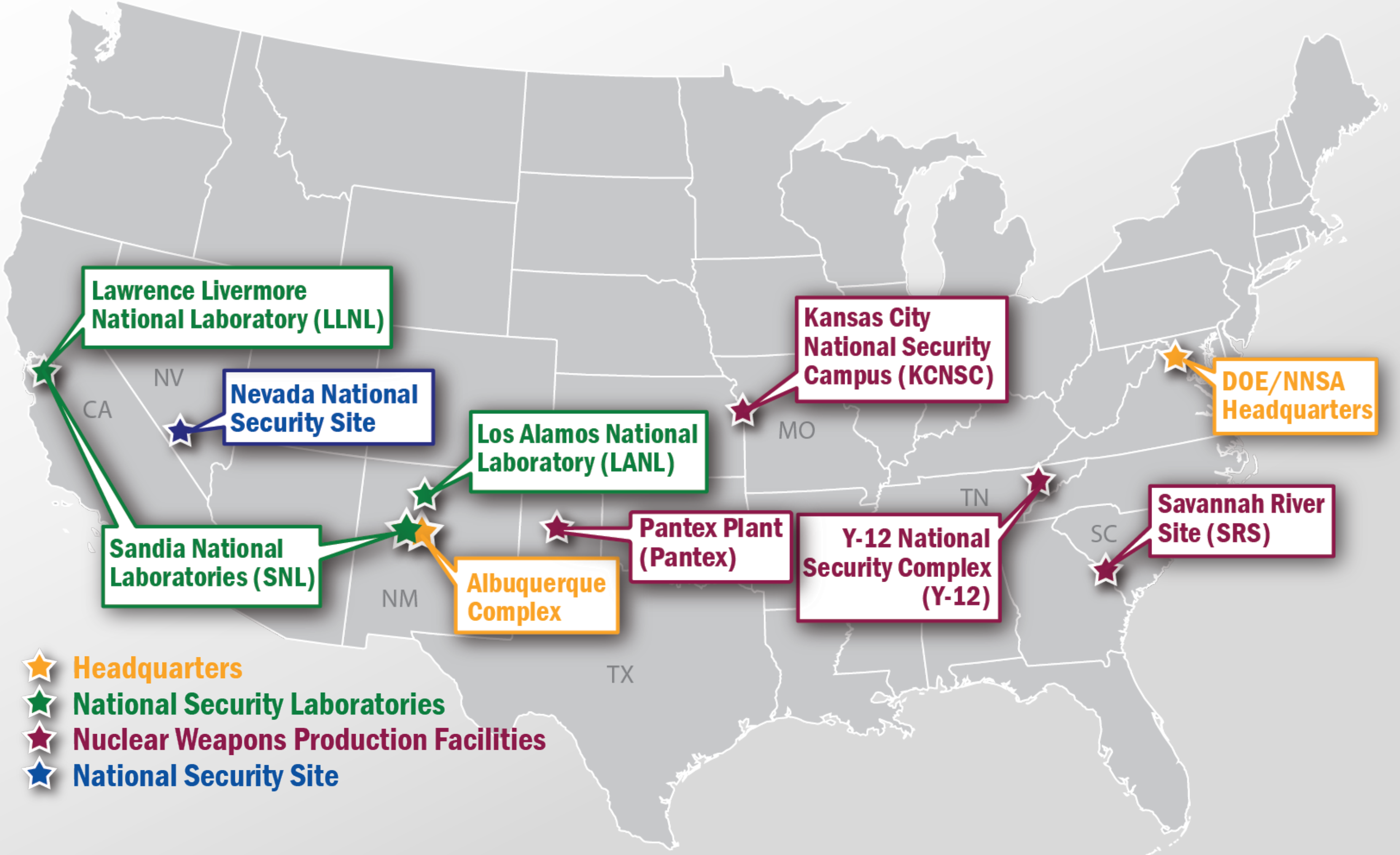
Reduce global **nuclear threats**

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.



Bottom Line Up Front

INNOVATE. COLLABORATE. DELIVER.

- **Purpose**: We can use Natural Language Processing to standardize different projects to a common WBS format
  - Maximizes the use of historical project cost data to develop early-stage estimates such as AOAs
  - Reduces the tendency of “reinventing the wheel” for each new analysis
- **Early work**: We demonstrated in an ICEAA 2022 paper/presentation how we used NLP to classify NNSA Capital Asset data to a high-level, six class scheme:
  - Site Preparation, Construction, Procurement, Project Management, and Start-up
- **Current work**: We developed a hierarchical NLP model that standardizes NNSA Capital Asset data to a multi-level, hierarchically-organized common WBS
  - Addresses known limitations of the initial, high-level model
  - Potential use-cases:
    - Standardize disparate projects’ cost data to a common WBS
    - Ability to leverage big data in cost and schedule estimating
    - Compare different WBS schemes based on how well they can classify historical cost data
    - Validate a project’s compliance to a known, agreed-upon standard

The Data

- PARS NNSA project cost data in the form of a WBS
  - Various projects, some more similar than others
    - **New construction, equipment install, refurbishment, D&D, combinations of these**
- Multi-input
  - WBS elements – Text, hierarchical
  - Cost data – Quantitative, includes negative values
  - Dates – Time series

- Multi-Output
  - Developed a standard WBS that goes up to 8 levels deep, depending on the branch

Previous Work – Model 1.0

INNOVATE. COLLABORATE. DELIVER.

- Multi-input model
- Maps cost data to a high-level (Level 2) WBS
- Limitations
  - Used text from the lowest level
  - Labeling
  - Limited samples
  - Biased dataset
  - Imbalanced Classes

Level	WBS Code	Title
1	1	Project Name
2	1.1	Site Preparation
2	1.2	PED
2	1.3	Construction
2	1.4	Procurement
2	1.5	Project Management
2	1.6	Start-up

Work Breakdown Structure

- Hierarchical, path-dependent, tree-like structure
- Contains varying levels, samples per level

Known Challenges:

- Prone to class-imbalance
- Importance of levels varies between projects
- Quality of text varies between projects
- Depth of the levels varies between paths
- Typical NLP-related challenges

General Classification Types

- Single Label v. Multi-Label
- Multioutput
- Multi-input Multioutput (MIMO)
- Hierarchical Classification

Mapping data to a high-level WBS is a single-label problem.

Mapping data to a multi-level WBS is a single-label, multi-output problem.

New Labeling Procedure

(a)

Project	WBS	Title
A	1	Project A
A	1.1	Equipment
A	1.1.1	Install
A	1.1.1.1	Glovebox
B1	A	Project B
B1	A.1	Subproject B1
B1	A.1.a	Procurement
B1	A.1.a.i	Equipment
B1	A.1.a.i.a	Glvbox

(b)

Model 1 Labeling Procedure using WBS

Labeled Element	Level 2	Level 3	Level 4	Level 5
Glovebox	Equipment	?	?	?
Glovebox	Equipment	?	?	?

(c)

Model 2 Labeling Procedure using WBS

Labeled Element	Level 2	Level 3	Level 4	Level 5
Equipment Install Glovebox	Equipment	Installation & Integration	Process Equipment	Installation
Procurement Equipment Glovebox	Equipment	Procurement	Process Equipment	--

Model 2 labelling procedure now adds the higher-level element text to their corresponding lower-level elements.

Augmenting the Data

INNOVATE. COLLABORATE. DELIVER.

Model	Project	Level 2	Level 3	Labeled Element	Unique Element
1.0	A	Equipment	Installation	GB	**GB**
	B	Procurement	Equipment	GB	
2.0	A	Equipment	Installation	Equipment Installation GB	**Equipment Installation GB**
	B	Procurement	Equipment	Procurement Equipment GB	**Procurement Equipment GB**

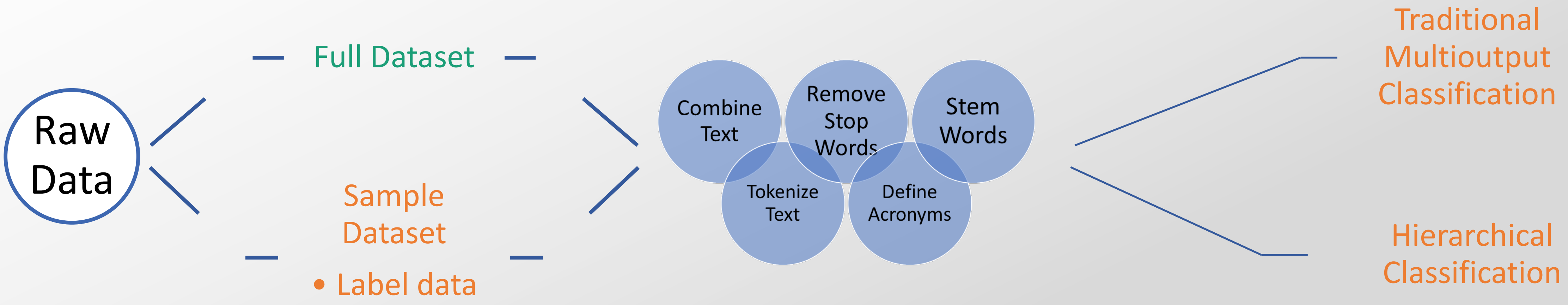
Model	Unique Element	Level 2	Level 3	Level 4	Level 5
1.0	**GB**	Equipment	Process Equipment	?	?
2.0	**Equipment Installation GB**	Equipment	Installation	Process Equipment	Installation
	Procurement Equipment GB	Equipment	Procurement	Process Equipment	--

Concatenating the text increases the number unique elements used for training and provides more text information for the model

NNSA National Nuclear Security Administration

Data Preprocessing

INNOVATE. COLLABORATE. DELIVER.



Current Work: WBS Model 2.0

INNOVATE. COLLABORATE. DELIVER.

Common Metrics	New Metrics
▪ Precision $= \frac{True\ Positive}{True\ Positive + False\ Positive}$	▪ Hierarchical Precision (hP) $= \frac{\sum_i \lvert a_i \cap B_i \rvert}{\sum_i \lvert a_i \rvert}$
▪ Recall $= \frac{True\ Positive}{True\ Positive + False\ Negative}$	▪ Hierarchical Recall (hR) $= \frac{\sum_i \lvert a_i \cap B_i \rvert}{\sum_i \lvert B_i \rvert}$
▪ F1-score $= 2\ x\ \frac{Precision * Recall}{Precision + Recall}$	▪ Hierarchical F1-score (hF) $= 2$ x $\frac{hP * hR}{hP + hR}$

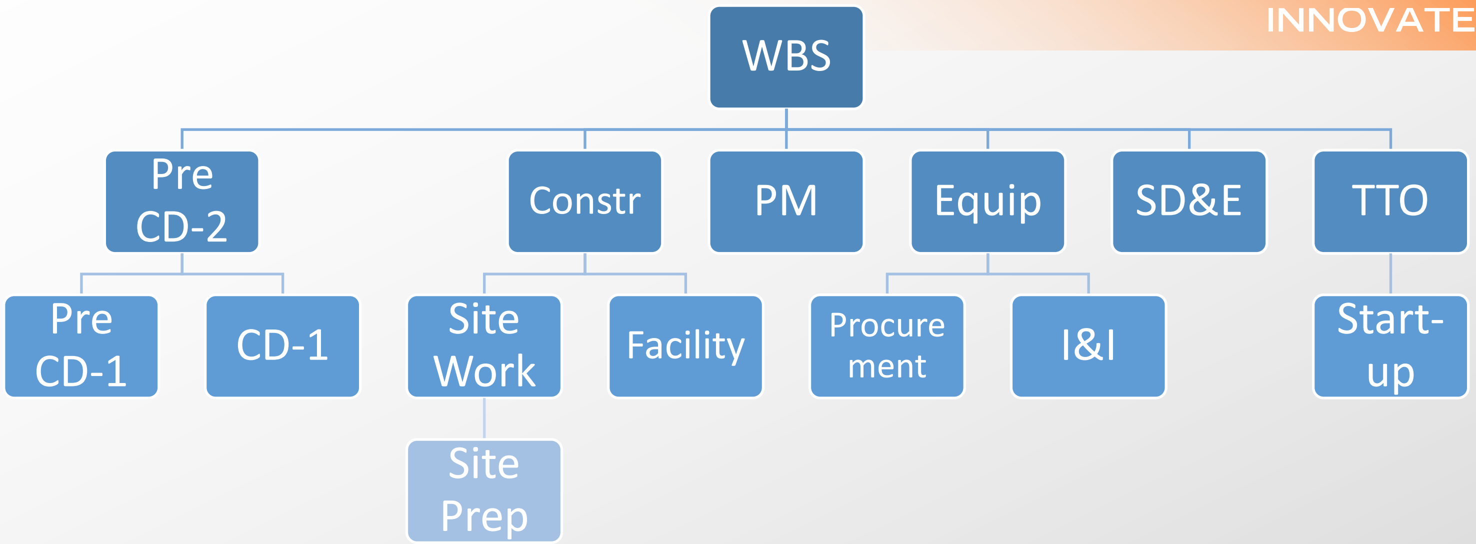
a_i = set with the most specific class *predicted* for test example *i* and its ancestor classes

B_i = set containing the *true* most specific class of test example *i* and all its ancestor classes

The hierarchical scores are based on summations across all test examples

National Nuclear Security Administration

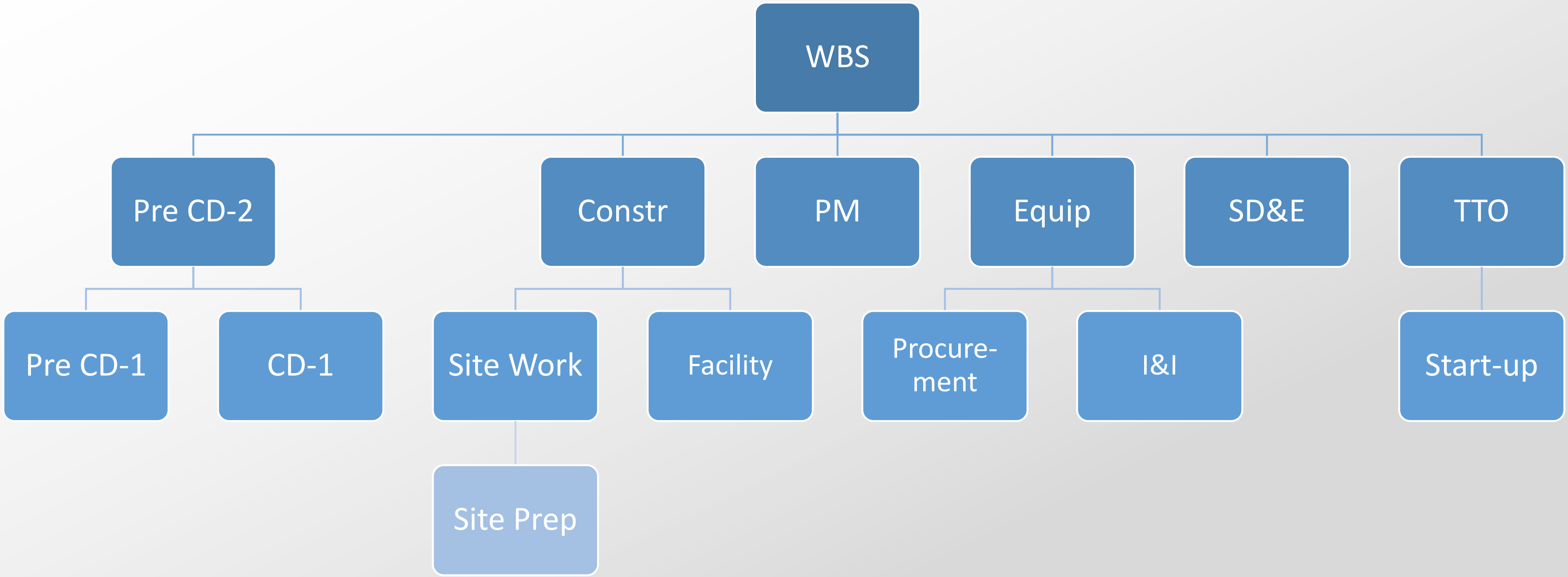
Example



Predicted	True	a_i	B_i	$a_i \bigcap B_i$	hP	hR	$\boldsymbol{hF}$
Pre CD-1	CD-1	2	2	1	1/2 = 0.5	1/1 = 1.0	.67
Site Work	CD-1	2	2	0	0/2 = 0	0/0 = NA	NA
CD-1	CD-1	2	2	2	2/2 = 1.0	2/2 = 1.0	1.0
TOTAL		6	6	3	3/6 = 0.5	3/3 = 1.0	.67

Simplified WBS

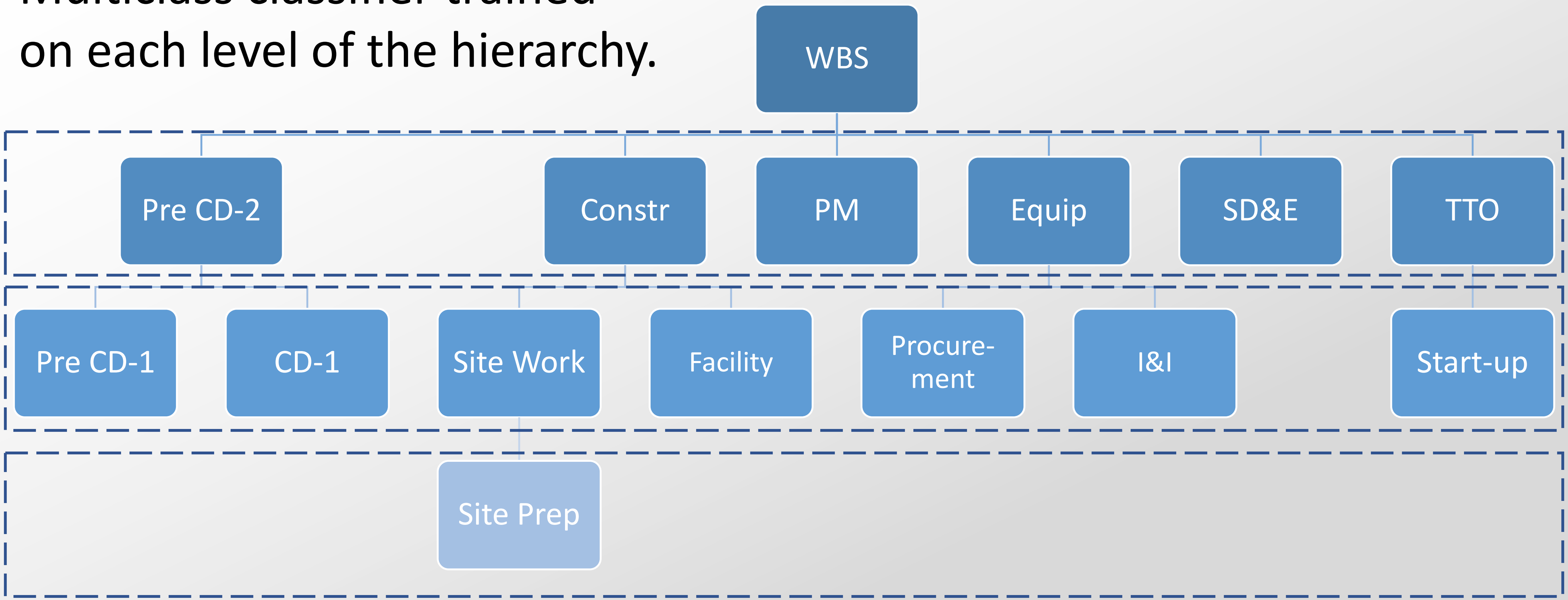
INNOVATE. COLLABORATE. DELIVER.



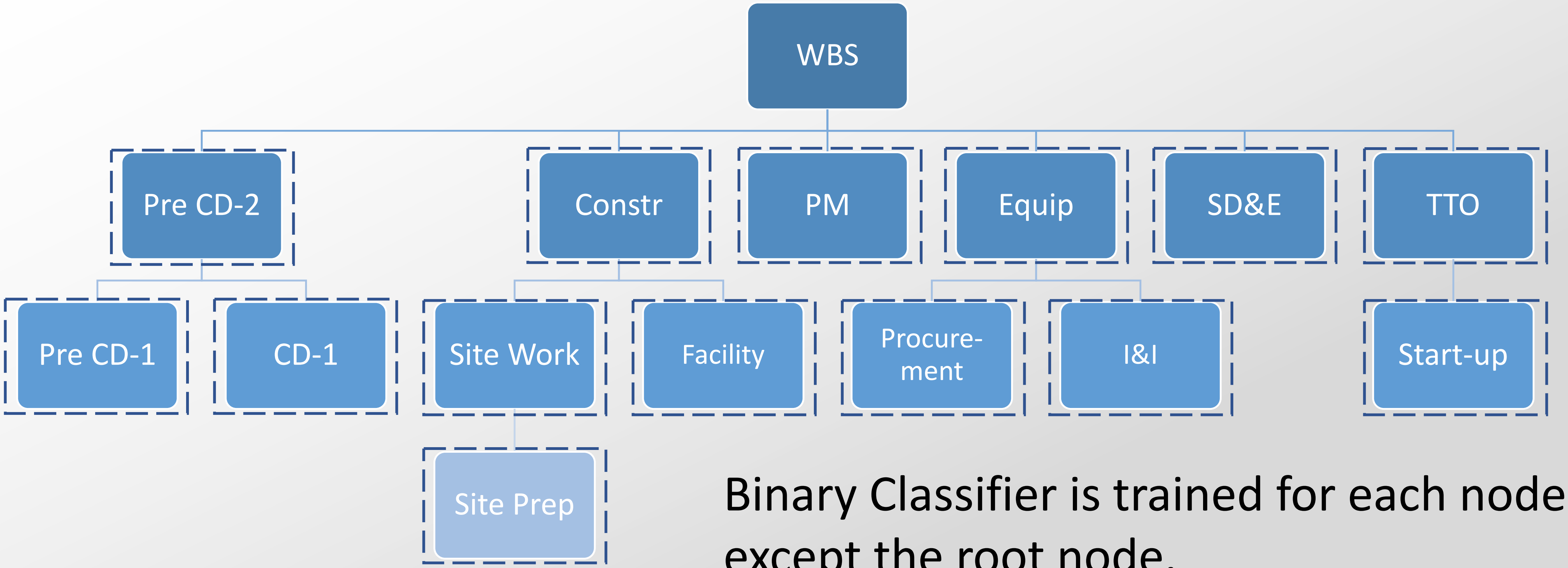
Local Classifier Per Level

INNOVATE. COLLABORATE. DELIVER.

Multiclass classifier trained on each level of the hierarchy.

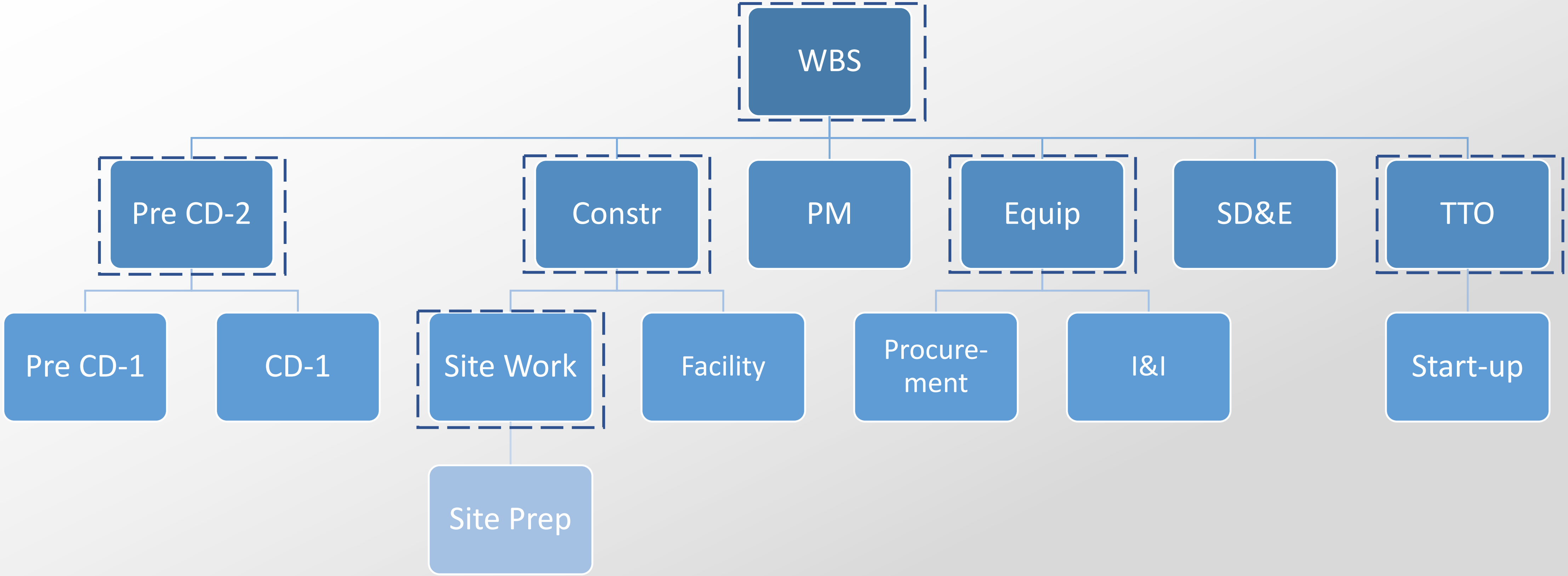


Local Classifier Per Node



Binary Classifier is trained for each node except the root node.

Local Classifier Per Parent Node



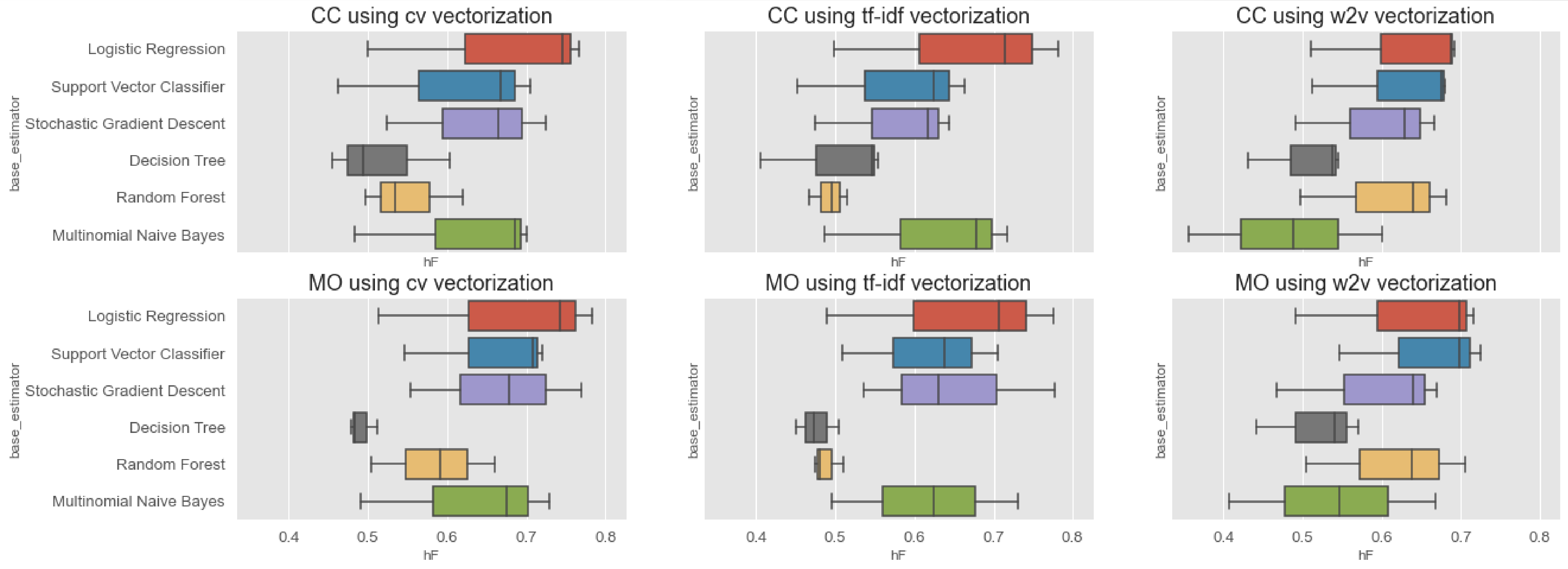
A Multiclass classifier trained for each parent node to predict the child nodes.

Model 2 Candidates

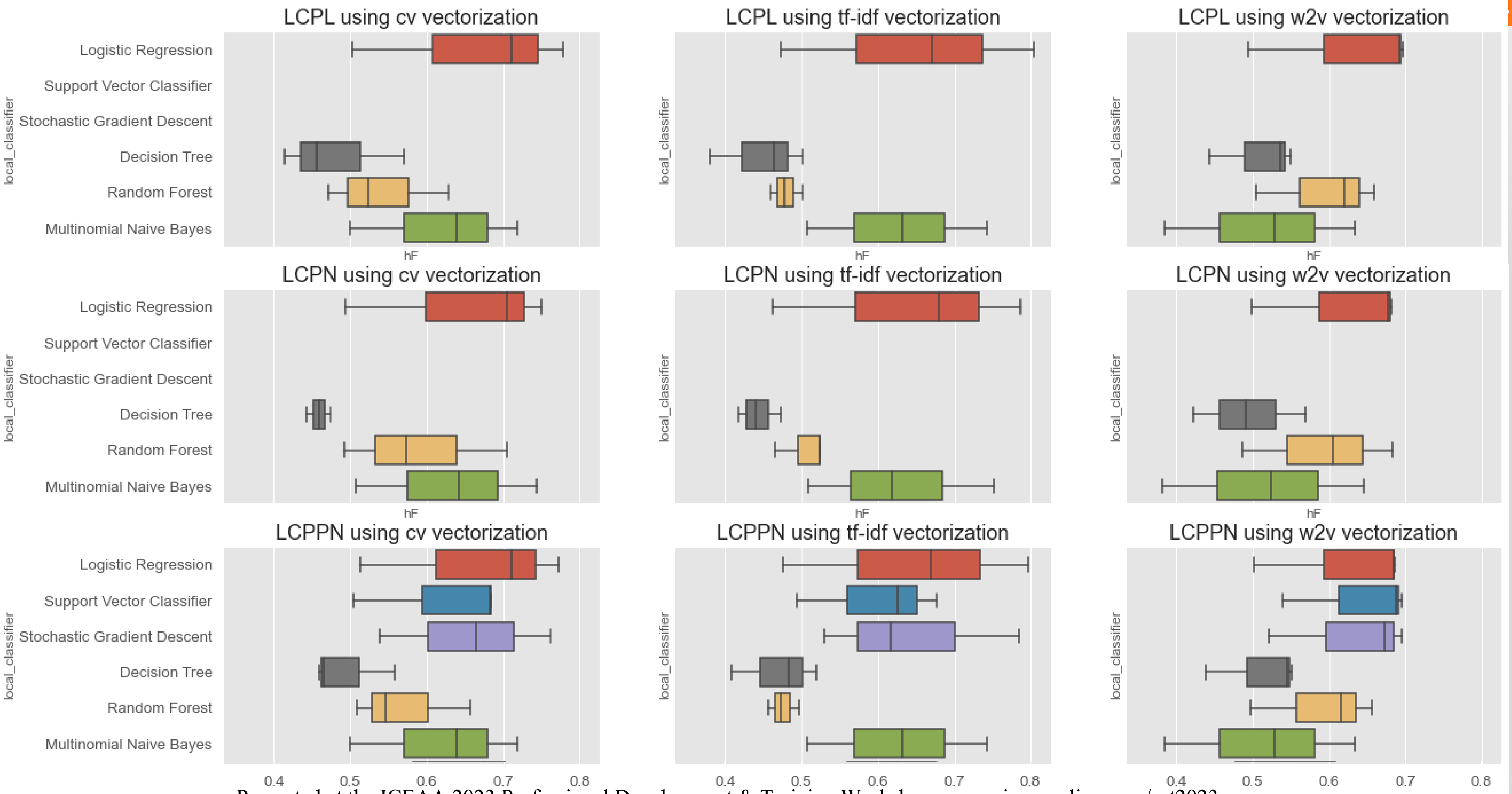
Meta-Estimators	Vectorizers	Base Estimators / Local Classifiers
• Multioutput Classifier • Classifier Chain • Local Classifier Per Level • Local Classifier Per Node • Local Classifier Per Parent Node	• Count vectorization • Tf-idf vectorization • Word2vec vectorization	• Support Vector Classifier • Stochastic Gradient Descent • Decision Tree • Multinomial Naïve Bayes • Random Forest • Logistic Regression

Initial results - Non-hierarchical Models

INNOVATE. COLLABORATE. DELIVER.



Initial results - Hierarchical Models



Model Results – Non-hierarchical Models

INNOVATE. COLLABORATE. DELIVER.

Meta-Estimator	Vectorizer	Best Base Estimator	hF
Multioutput	count	Stochastic Gradient Descent	0.875
	tf-idf	Stochastic Gradient Descent	0.874
	word2vec	Support Vector Classifier	0.831
Classifier Chain	cv	Stochastic Gradient Descent	0.777
	tf-idf	Stochastic Gradient Descent	0.782
	word2vec	Support Vector Classifier	0.811

Model Results – Non-hierarchical Models

Meta-Estimator	Vectorizer	Best Base Estimator	hF
Local Classifier Per Parent Node	count	Stochastic Gradient Descent	0.889
	tf-idf	Stochastic Gradient Descent	0.809
	word2vec	Stochastic Gradient Descent	0.792

Model Results – Untuned Neural Nets

INNOVATE. COLLABORATE. DELIVER.

Meta-Estimator	Vectorizer	Base Estimator	Best hF
Multioutput	int	Dense Neural Net	0.563
	int	Convolutional Neural Net	0.648
Classifier Chain	int	Dense Neural Net	--
	int	Convolutional Neural Net	-

Model Results – Untuned Neural Nets

Meta-Estimator	Vectorizer	Local Classifier	Best hF
Local Classifier Per Parent Node	int	Dense Neural Net	0.643
	int	Convolutional Neural Net	0.639

Model Results – Tuned Neural Nets

Meta-Estimator	Base Estimator	Best Hyperparameter Values	hF
Multioutput	Dense Neural Net	Embedding Dimension: 100 Hidden Layers: 2 Dropout: .5 Neurons per Layer: 200	--
	Convolutional Neural Net	Embedding Dimension: 200 Hidden Layers: 1 Number of Filters: 32 Number of Kernels: 10 Dropout: .3	--
Classifier Chain	Convolutional Neural Net	--	--
	Dense Neural Net	--	--

Model Results – Tuned Neural Nets

INNOVATE. COLLABORATE. DELIVER.

Meta-Estimator	Base Estimator	Best Hyperparameter Values	hF
Local Classifier Per Parent Node	Dense Neural Net	Embedding Dimension: 200 Hidden Layers: 1 Dropout: .3 Neurons per Layer: 200	0.854
	Convolutional Neural Net	Embedding Dimension: 200 Hidden Layers: 2 Number of Filters: 64 Number of Kernels: 15 Dropout: .5	0.896

Now that we have a functional and accurate model –

What can we do with it?

Use Cases

- ***Time Savings***
  - **Status Quo: manual categorization of individual WBS'**
    - *Extremely time-intensive*
      - ~ 1 year to manually categorize 10 projects and create the model
    - *Interpretations vary between analysts*
  - **Now: NLP categorization of all available WBS'**
    - *Major time savings*
      - Model categorized the remaining 137 projects in the dataset + ability to automatically categorize all future projects added
    - *Analysts can focus on analyzing vs. wrangling data*
    - *Ability to provide more data-driven results and respond to quick-turn actions*

Use Cases

INNOVATE. COLLABORATE. DELIVER.

- ***Immediate Benefits:***
  - **Benchmarking costs for ongoing projects**
    - *"Is my project spending too much on oversight in the design phase, compared to other similar projects?"*
    - *"How much money do I need to budget for the preliminary design phase?"*

Project	PM costs as a % of A/E Design Costs
Project A	71%
Project B	23%
Project C	62%
Project D	181%
Project E	46%
Project F	76%
My Project	**128%**

Level 2	**Systems Design and Engineering**	**$19M**
Level 3	Title I/CD-2/Preliminary Design and Title II/CD-3 Final Design	$9M
Level 4	Title I / CD-2 Preliminary Design	$2.7M
Original Text	M&O Contractor Title I Design Support OPC	$.1M
Original Text	Title I AE Design Contract	$2.6M

Use Cases

- ***Immediate Benefits:***
  - **Leverage historic data for early-stage cost estimates**
    - More projects available to pick from for analogy estimating
    - Increase confidence in bounds on distributions for monte carlo simulations
    - Provide a large enough sample to allow for parametric modeling
    - Example:
      - Estimate cost to install equipment in a machine shop
        - Status quo: SME input, direct analogy from time spent categorizing a single previous project
        - NLP model: immediately able to choose analogies from any number of the 67 projects that installed equipment, build parametric CER to find drivers of installation cost, etc.

Use Cases

- ***Immediate Benefits:***
  - **Provides an exhaustive WBS structure and dictionary that can be leveraged for cost estimates and ongoing capital projects**
    - Enables internal consistency to compare two or more estimates that originally used different WBS formats
    - Dictionary provides definitions, examples, and lessons learned to allow other analysts to be able to interpret terminology on WBS’ and to understand what/when/why to pull information out of the model results

Use Cases

- ***Future Goals:***
  - **Consolidate disparate WBS formats and dictionaries into a standard**
    - Current procedure for new projects in DOE is to create a WBS format and dictionary for every new project
      - Efficiency loss within projects and information loss over time
    - Need buy-in and consensus from DOE stakeholders to adopt a standard for future projects
      - Until then, the model can serve as a translator and converter from a custom-built WBS to the model's standard

Use Cases

- ***Future Goals:***
  - **Apply lessons learned from model development to other data sets**
    - Greater ability to conduct machine learning elsewhere
    - Many organizations, both within and external to DOE, have large amounts of data, but without a full understanding of it
      - This model can be retrained on other datasets, providing insights much faster than building a new machine learning model from scratch

Post-Processing

- Management Plan - As additional requests for the model results come in, finding additional areas in the data to refine to improve our labeling
  - Retune whenever we get major new information
- Machine learning is not a perfect process, and there is some remaining work to be done to improve the overall accuracy of the data set

A Note on Alternatives

- Manual Approach
  - Reviews all data
  - Pros
    - Benefits from human intuition and the use of outside data sources to validate
  - Cons
    - Time consuming
    - Human error/inconsistency possible
    - Expensive in the long-term
- Deterministic Approach
  - Pros
    - Guaranteed results for exact matches
    - No model training required
  - Cons
    - Little flexibility for variation in text – need to anticipate deviations
    - No measures of performance

Conclusion

- The cost estimating and analysis community continues to receive more and more data each year.
- The community *must* adapt new methodologies to handle this data
- Hierarchical classification performed by utilizing NLP and machine learning is an emerging field and we look forward to continuing this work to further improve our results and its utility to other use-cases within the cost estimating community

References

Géron, Aurélien. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow: Concepts, tools, and techniques to build intelligent systems.* O'Reilly Media.

Badaracco, A. G. (2022). *SciKeras 0.9.0 Documentation.* Retrieved from SciKeras: https://www.adriangb.com/scikeras/stable/index.html

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.

Miranda, F. M., Köehnecke, N., & Renard, B. Y. (2021). HiClass: a Python library for local hierarchical classification compatible with scikit-learn. *arXiv preprint arXiv:2112.06560* .

Sechidis, K., Tsoumakas, G., & Vlahavas, I. (2011). On the Stratification of Multi-Label Data. *Machine Learning and Knowledge Discovery in Databases*, 145-158.

Silla, C., & Freitas, A. (2011). A survey of hierarchical classification across different application domains. Data Mining and Knowledge Discovery, 22, 31-72. Retrieved from https://doi.org/10.1007/s10618-010-0175-9

Szymański, P., & Kajdanowicz, T. (2017). A Network Perspective on Stratification of Multi-Label Data. *Proceedings of Machine Learning Research* (pp. 22-35). ECML-PKDD, Skopje, Macedonia: PMLR.

Thank you!

mlapoff@technomics.net
pshelton@technomics.net