EKE-Electronics Ltd.

IRIS® Certification

Designing Effort Estimation Process for Embedded Software Projects

Master's thesis presentation
by Amir Sultanbekov

23/09/2024

Thesis supervisor
Professor Ahti Salo

Thesis advisor
Juha Paldanius

Smarter trains. Better future.

EKE

Contents

- EKE-Electronics Ltd.
- Effort estimation challenges
- Proposed approach
  - Overview
  - Method selection
  - Expert-based estimation
  - Machine learning models
  - Implementation plan

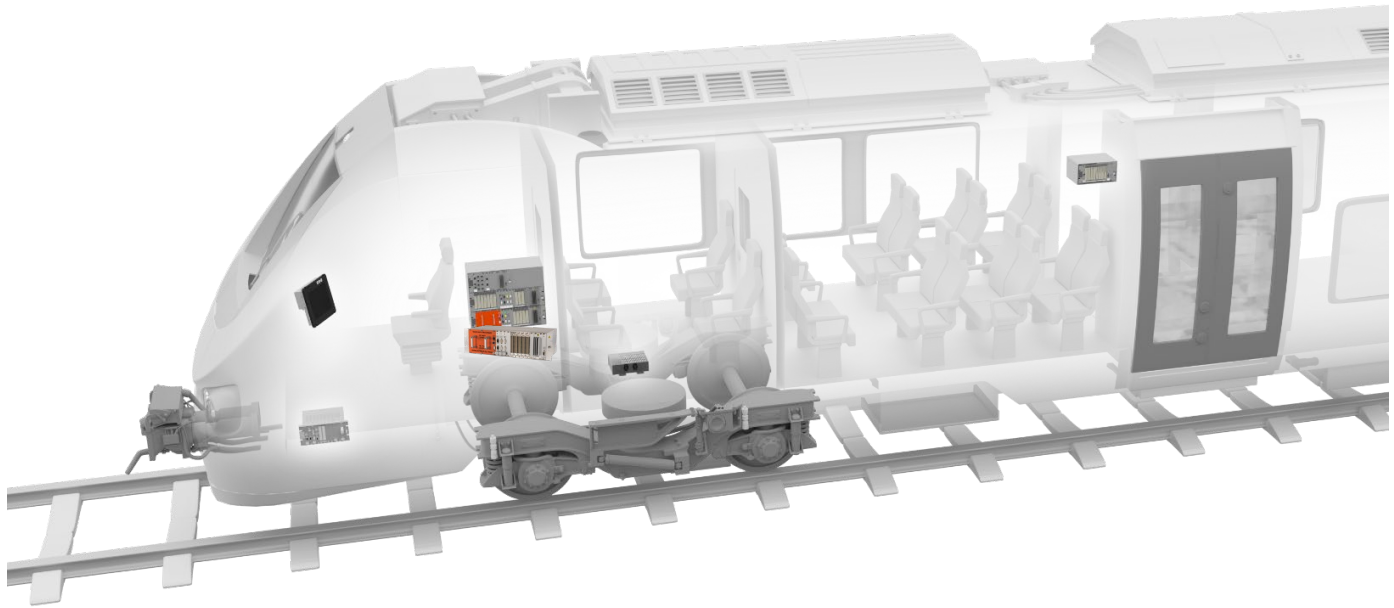
EKE-Electronics

1986
Established

100
Employees

35 000+
Train automation systems

6
Continents



Smarter trains. Better future.

EKE-Electronics

1986
Established

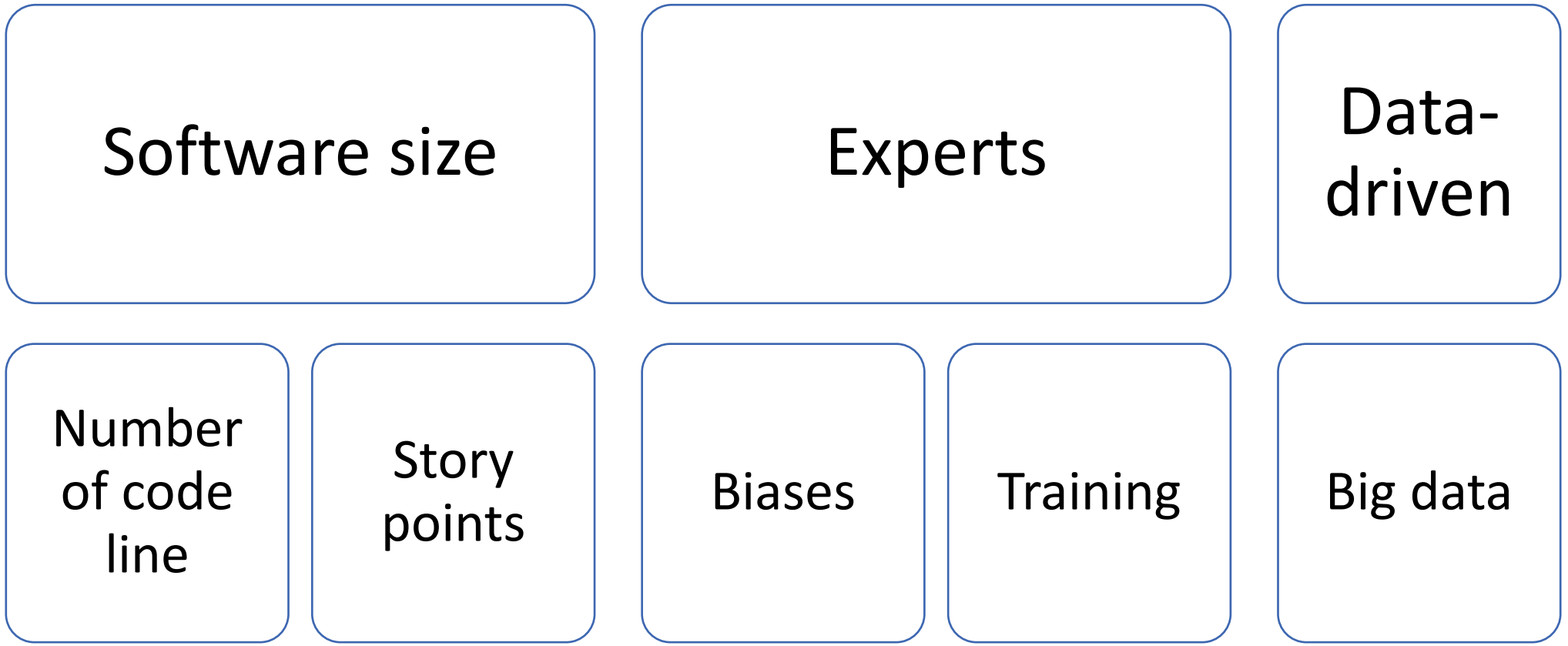
100
Employees

35 000+
Train automation systems

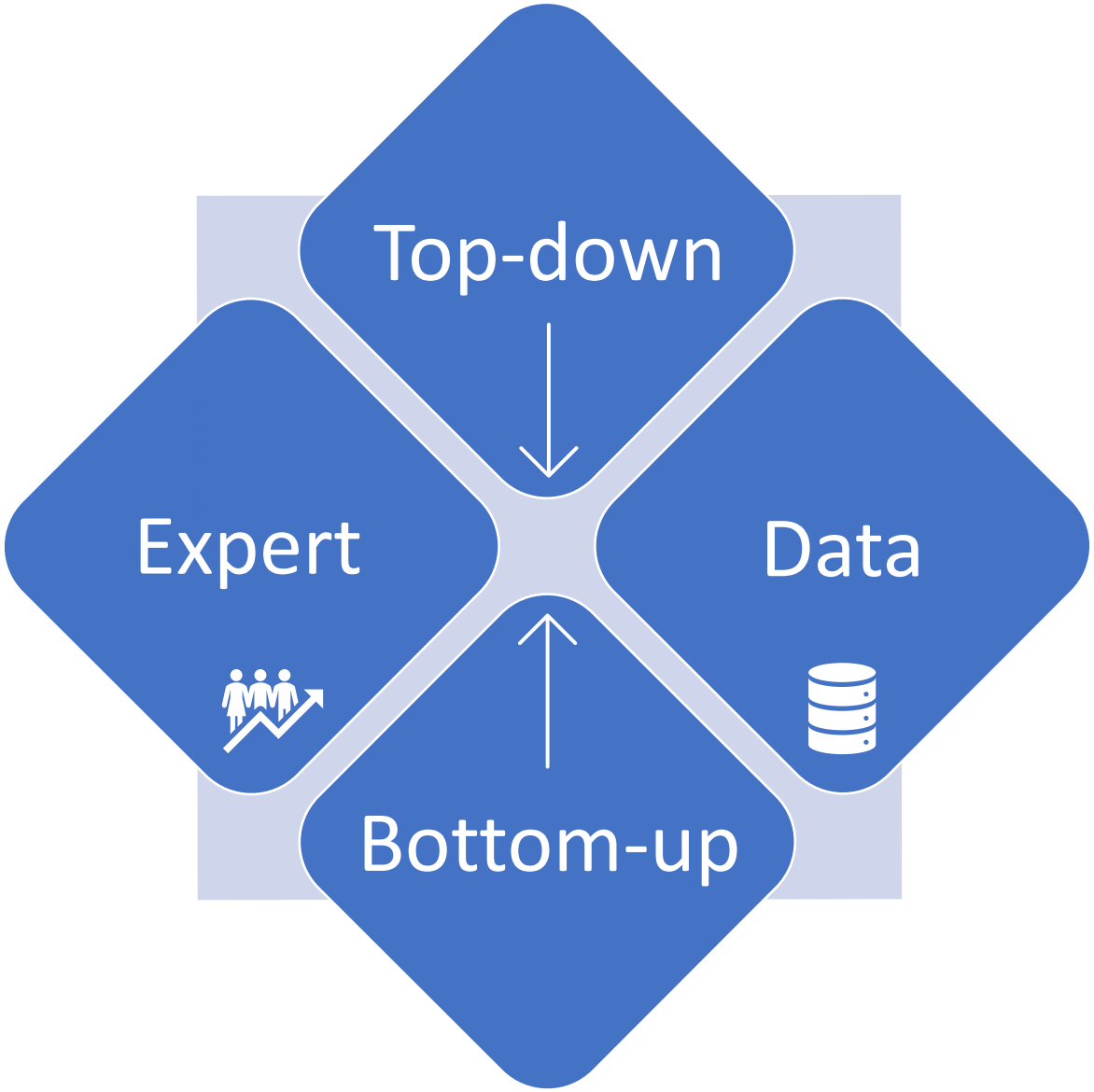
6
Continents

Smarter trains. Better future.

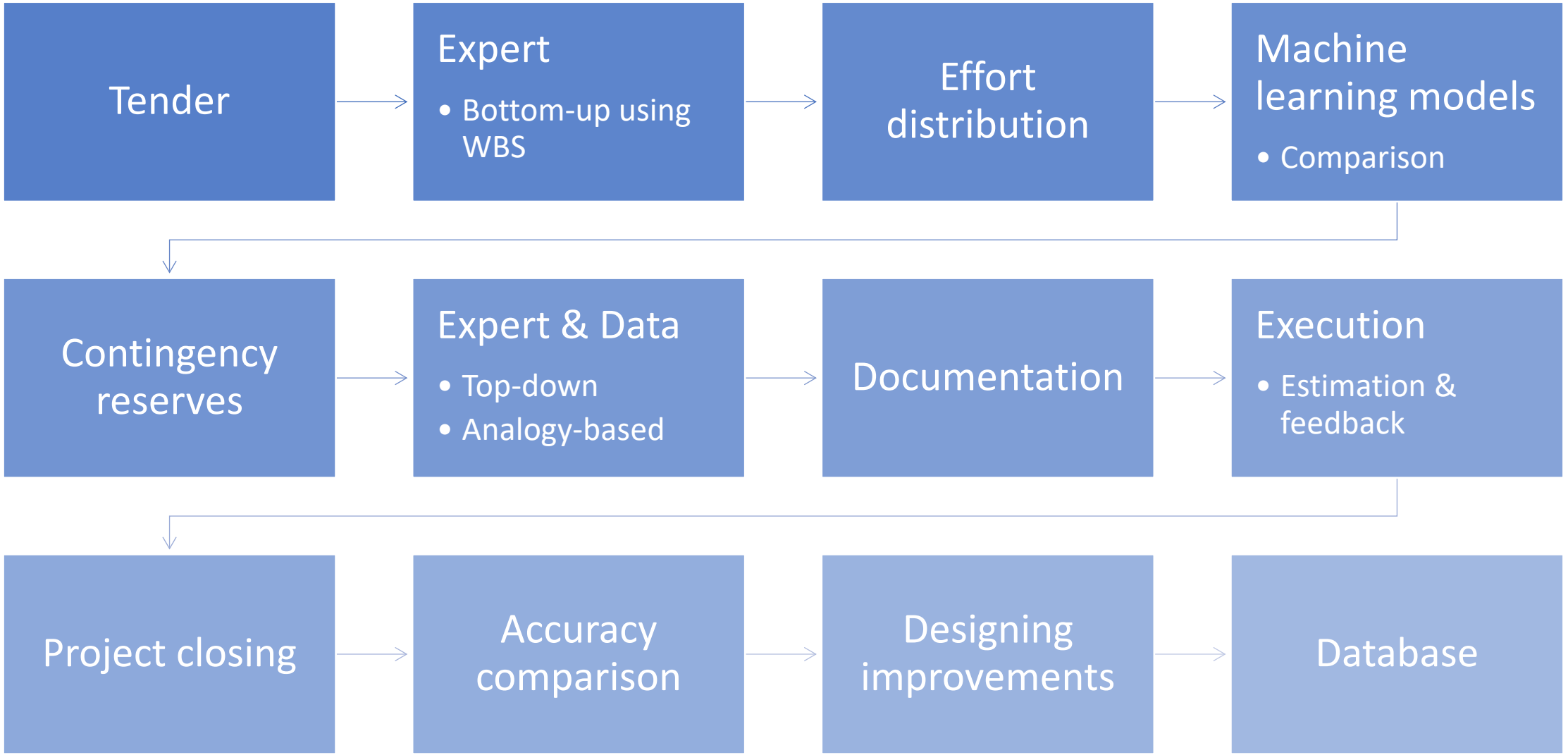
Estimation challenges



Proposed approach



Process overview



Method selection

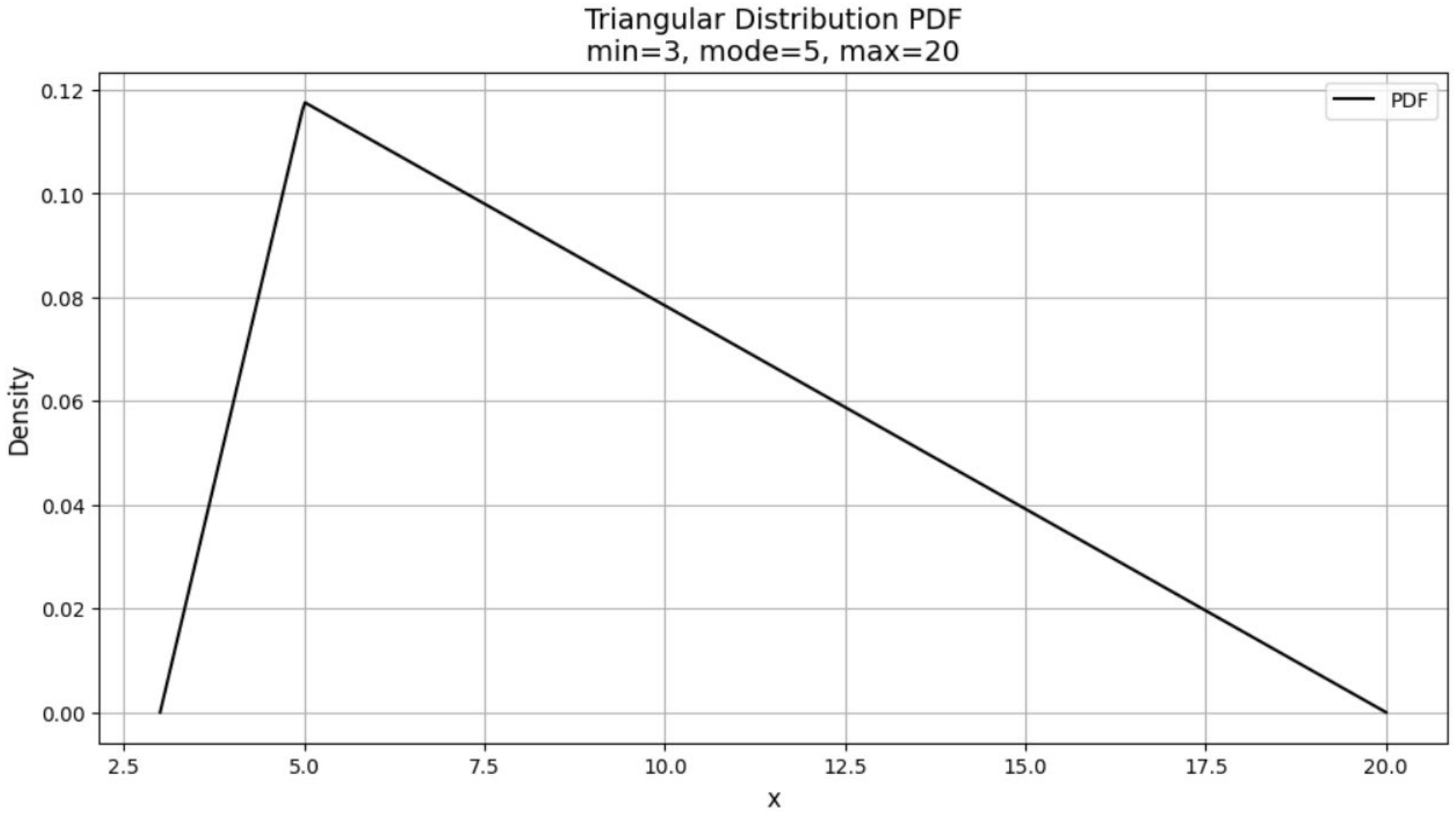
Criterion	Expert involvement	Required data	Robustness	Flexibility	Complexity	Support	Reusability	Predictive power	Informative power	Handling uncertainty	Comprehensiveness	Empirical evidence
Expert involvement		2	2	3	5	6	4			5	8	7
Required data			2	2	5	5	3			4	7	6
Robustness				2	4	5	2			3	6	5
Flexibility					3	4	2			2	5	5
Complexity						2					3	2
Support											2	2
Reusability					2	3				2	5	4
Predictive power	2	3	4	5	7	8	5		2	6	9	8
Informative power	2	2	3	4	6	7	5			5	8	8
Handling uncertainty					2	2					4	3
Comprehensiveness												
Empirical evidence											2	

Criterion	Normalized weight	Classical Model	IDEA	Wideband Delphi	Planning Poker	CBR	Regression	Random forest	Catboost / XGBoost	ANN	COCOMO
Expert involvement	0,15	3	2	1	1	4	5	5	5	5	4
Required data	0,11	4	4	5	5	1	1	2	2	2	4
Robustness	0,08	3	3	2	2	4	1	3	4	4	3
Flexibility	0,06	5	5	5	5	4	3	4	4	4	1
Complexity	0,03	2	2	5	5	4	4	3	2	2	3
Support	0,02	4	4	5	5	4	5	5	4	4	5
Reusability	0,05	1	1	1	2	3	4	4	4	4	5
Predictive power	0,24	3	4	2	2	4	1	2	4	4	1
Informative power	0,19	4	4	2	2	3	5	4	3	3	2
Handling uncertainty	0,04	5	5	3	3	2	1	1	1	1	1
Comprehensiveness	0,02	5	5	5	5	4	4	4	4	4	1
Empirical evidence	0,02	2	2	5	5	5	3	4	2	2	5
Sum	1	3,4	3,5	2,6	2,7	3,4	2,9	3,2	3,5	3,5	2,5

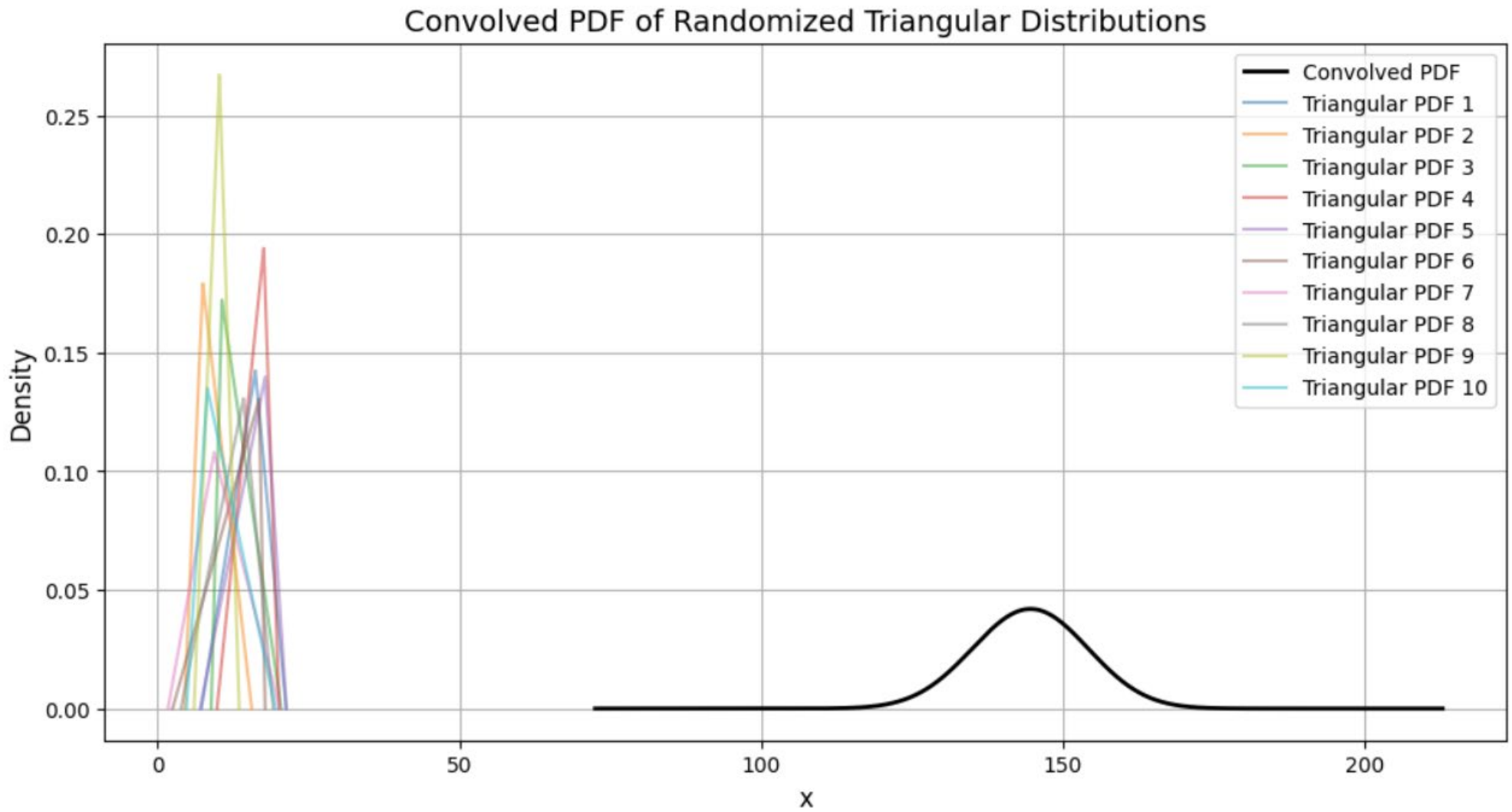
Expert-based bottom-up estimation

Cat A	Cat B	Cat C
2-3 engineers	2 engineers	1 engineer
Lowest plausible value?		
Highest plausible value?		
Likely value?		
How confident are you?		

Data distributions

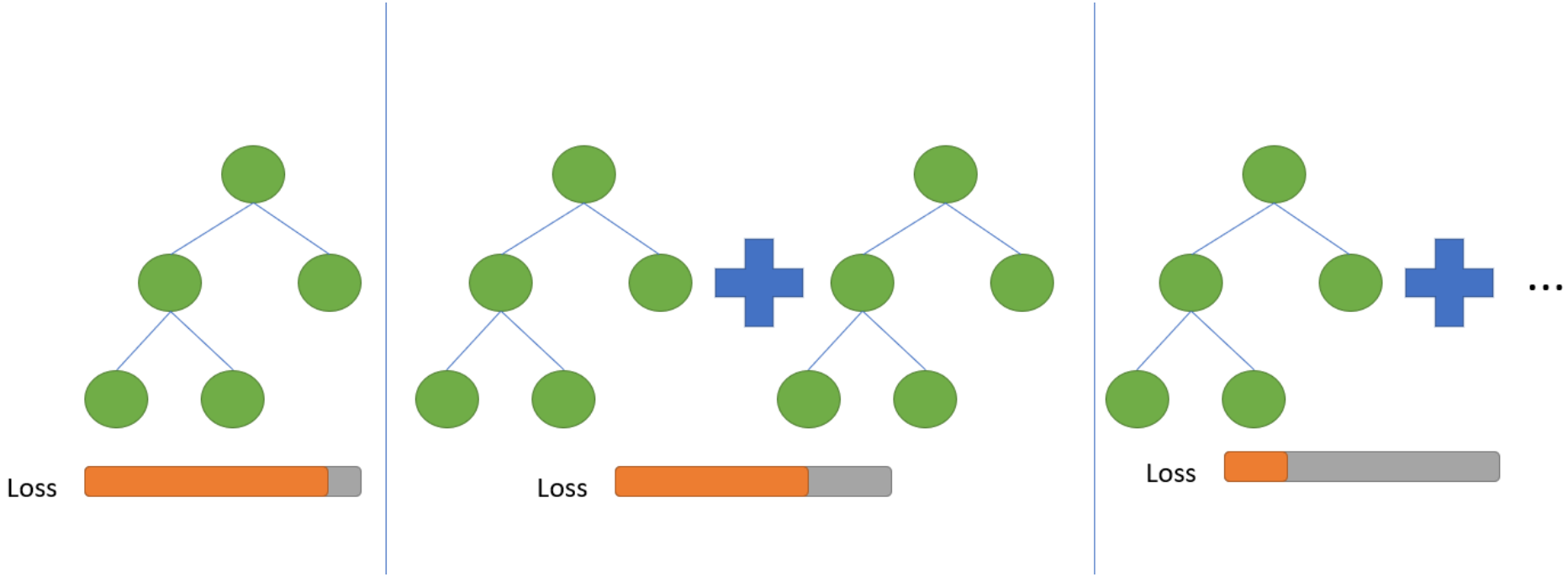


Data distributions



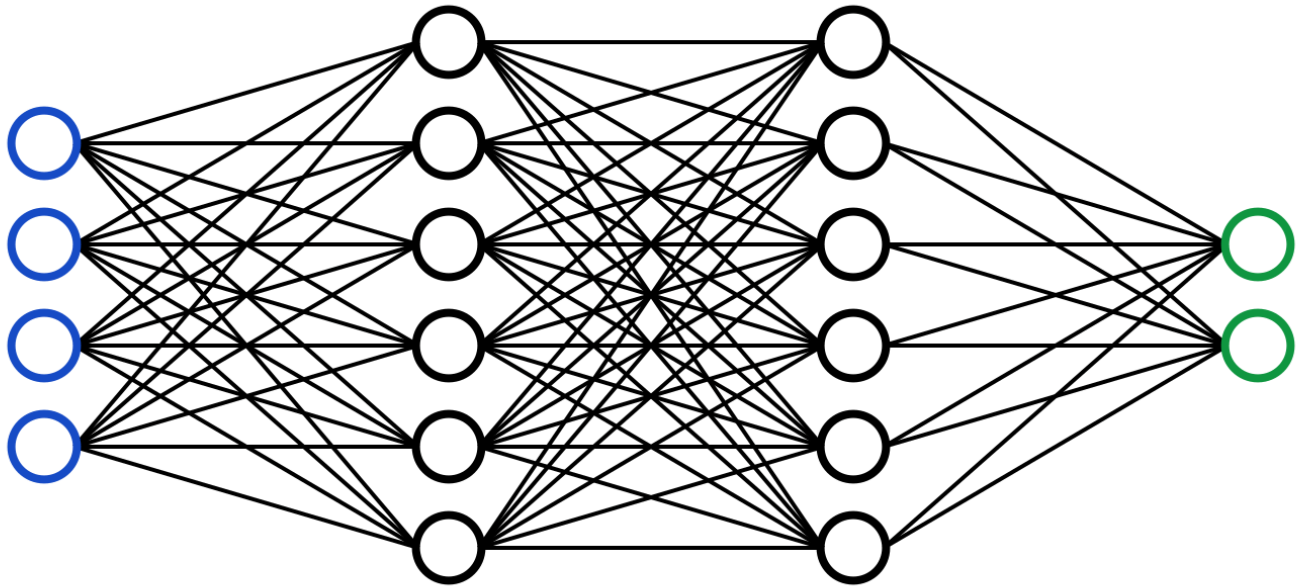
Machine learning models

- CatBoost, Extreme Gradient Boosting (XGBoost), artificial neural network
- Automatic tuning with Optuna
- Model structure explanation using Shapley Additive Explanations (SHAP)



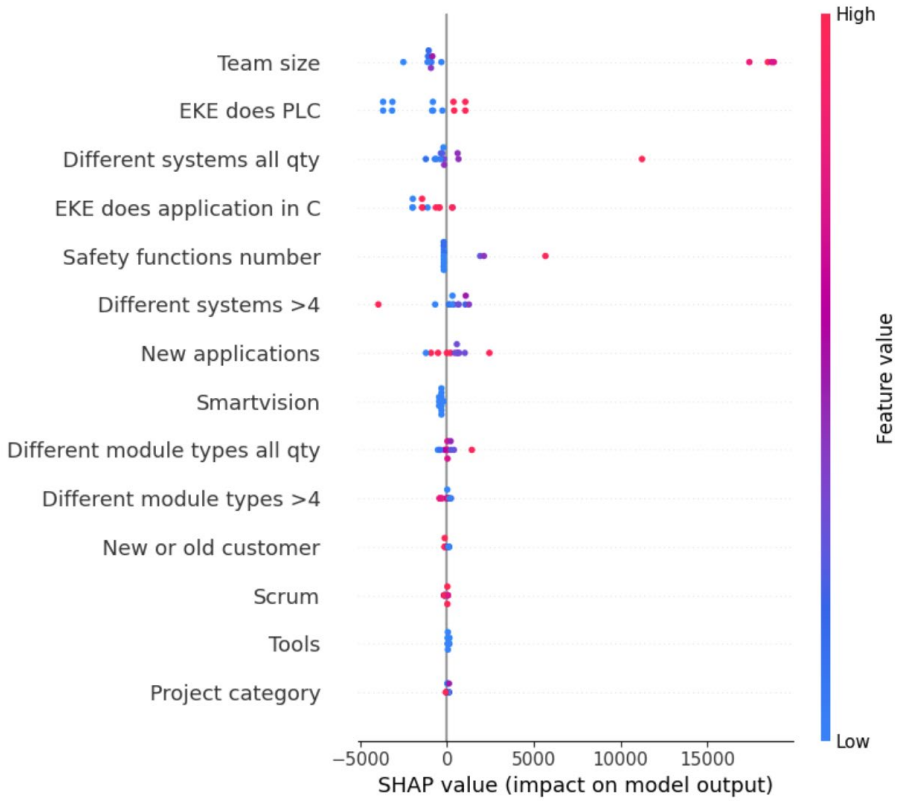
Machine learning models

- CatBoost, Extreme Gradient Boosting (XGBoost), artificial neural network
- Automatic tuning with Optuna
- Model structure explanation using Shapley Additive Explanations (SHAP)



Machine learning models

- CatBoost, Extreme Gradient Boosting (XGBoost), artificial neural network
- Automatic tuning with Optuna
- Model structure explanation using Shapley Additive Explanations (SHAP)



Machine learning results & Process implementation

Metric	CatBoost	XGBoost	ANN	Current
MSE	3*10^7	2*10^7	10^8	
MSBRE	8	7	1	
PRED(25%)	17 %	17 %	17 %	20 %
PRED(50%)	67 %	50 %	50 %	30 %

- Kotter's 8-stage process