

TABLE IX: Evaluation of Different NLP Approaches

Research	NLP	Input	Output	Dataset	Approach	Tool/Library	Input Structure	Evaluation Criteria
[29]	Tokenization POS Conference resolution Stemming Type dependencies	User story	Class diagram	Inline course management	Heuristic	Stanford CoreNLP WordNet PyEcore API PlantUML API Visual Narrator	Semi-structured	Performance
[27]	Tokenization POS Stemming Lemmatization Type dependencies	English text	Usecase diagram Activity diagram	Qualification Verification System [2]	Heuristic	Stanford CoreNLP	Unstructured	Applicability
[41]	Tokenization POS Splitting Syntax tree Chunking	English text	Class diagram OCL	-	Heuristic	NLTK	Unstructured	Precision Recall Applicability
[20]	Insufficient information	English text	Usecase diagram Activity diagram	-	Model transformation	Stanford POS Tagger	Unstructured	Applicability
[1]	Tokenization POS	Scenario-based requirement	Sequence diagram Class diagram	-	Heuristic	Stanford CoreNLP PlantUML	Semi-structured	Accuracy
[8]	Stop-word-removal Tokenization POS	English text	Class diagram	Automated mobile Money transfer System [4]	Heuristic	-	Unstructured	Applicability
[14]	Stemming Lemmatization POS Type dependencies	User story	Usecase diagram	User story [7]	Heuristic	-	Semi-structured	Syntactic correctness Semantic correctness Applicability
[34]	Stop-word-removal Tokenization POS	English text	Class diagram	PIMS [18] CR [17] WAAI [69]	Heuristic	Stanford Core NLP	Semi-structured	Precision Recall
[38]	Stop-word-removal Tokenization POS Type dependency Chunking	English text	Class diagram	ATM [67] LMS [68] OS [30]	Heuristic	SpaCy Windows Presentation Foundation(WPF)	Unstructured	Recall Precision Recall
[15]	Tokenization POS	User story	Goal model	-	Heuristic	-	Semi-structured	-

TABLE X: Evaluation of Different ML/DL Approaches

Research	NLP	ML	Input	Output	Dataset/Case Study	Tool	Input Structure	Evaluation Criteria
[12]	Cleaning Tokenizing POS Type Dependency	Support Vector Machine Naive Bayes	English text	Class diagram	Public Requirement dataset [13]	-	Unstructured	Precision Recall F-measure
[28]	Tokenization POS Chunking Splitting	Weka vote algorithm Logistics and SMO	English text	Usecase diagram Class diagram	-	Visual studio Weka	Unstructured	Accuracy Applicability
[21]	Tokenization Stop-word-removal	Supervised learning	English text	ER diagram	-	-	Unstructured	Recall Precision
[2]	Tokenization Type dependencies Splitting POS	BPNN	English text	Usecase diagram	ATM System Cafeteria Ordering System (COS) library system	GATE ANNIE Stanford- CoreNLP	Unstructured	Precision Recall
[35]	Sentence splitting Tokenization Stemming POS tagging Coreference resolution Semantic word similarity Parse Tree	Insufficient information	User story	UML diagrams	Ticket Sale System	NLA WordNet ROMA Stanford Parser	Structured/ Unstructured	Recall Precision Fmeasure
[12]	Cleaning Tokenizing POS Type Dependency	Support Vector Machine Naive Bayes	English text	Class diagram	Public Requirement dataset [13]	-	Unstructured	Precision Recall F-measure