

HEICON

Global Engineering



KI-basierte Generierung von Testfalldefinitionen aus natürlichsprachigen Requirements



HEICON – Global Engineering GmbH

www.heicon-ulm.de | blog.heicon-ulm.de



Agenda

- ☞ Motivation
- ☞ Master's Thesis
 - ☞ Setup & Data
 - ☞ Parameter Name Detection
 - ☞ Parameter Value Detection
- ☞ Summary
- ☞ Future Trends: Modern Requirements



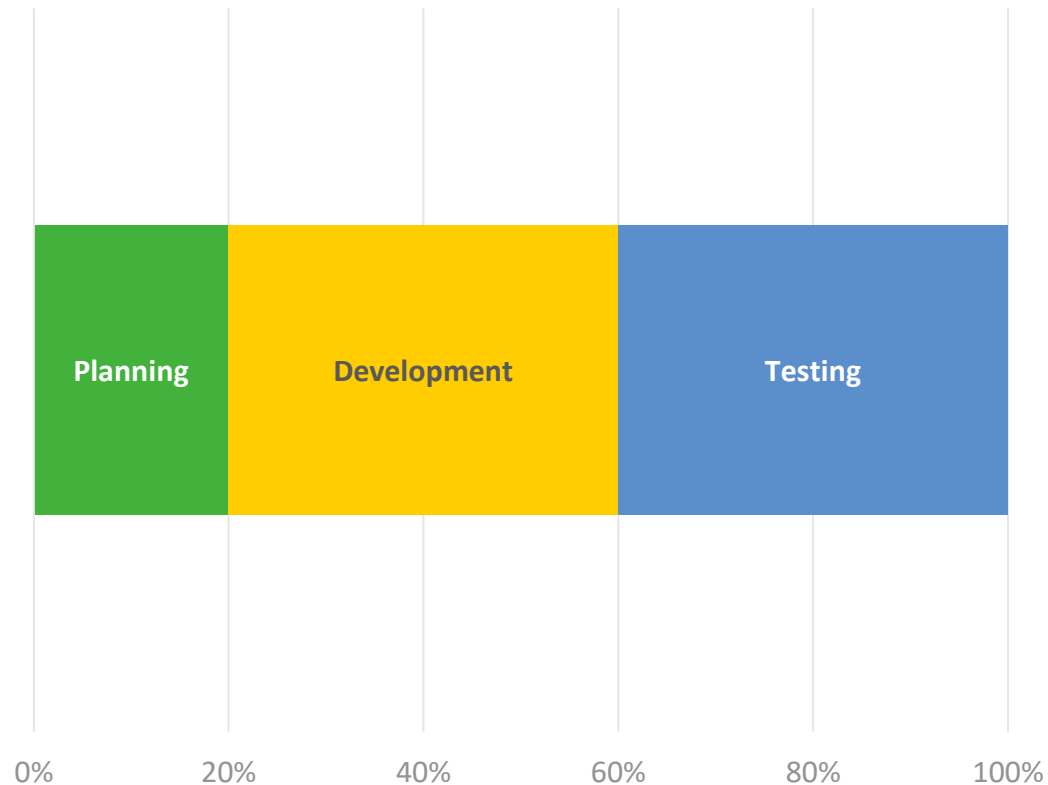
Motivation

AI supported test generation



The current state ...

Software project development efforts

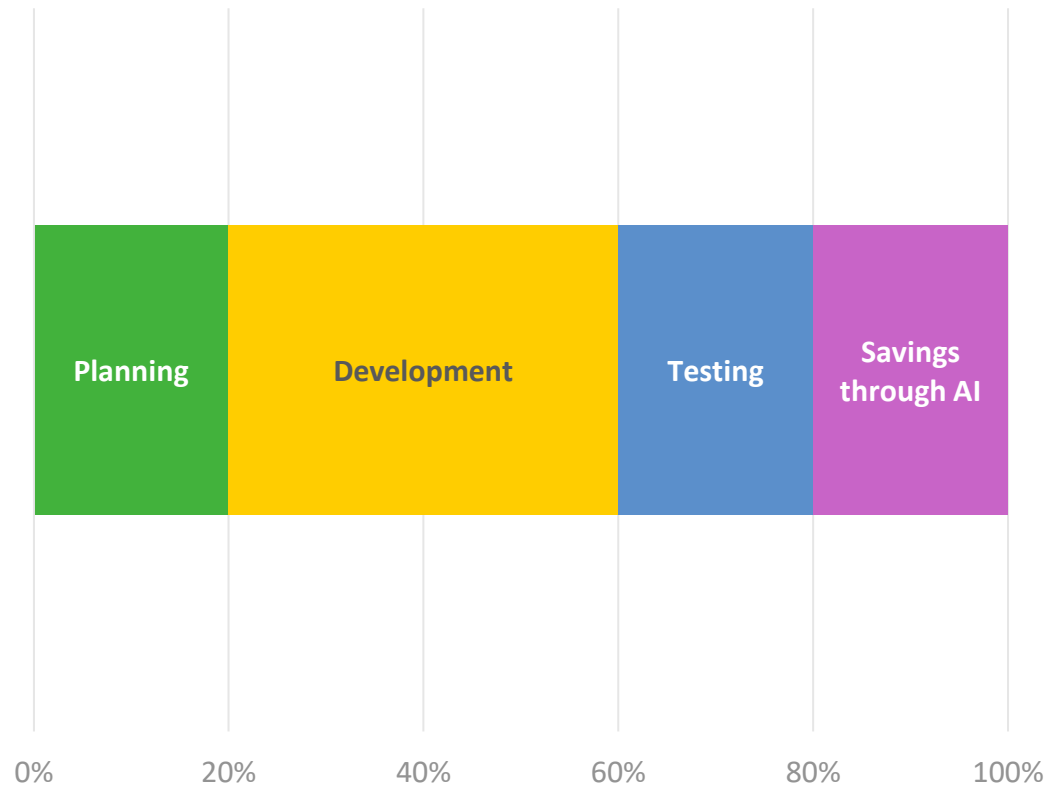


- Much more effort is put into testing, then into requirement engineering
- Test cases are derived manually from requirements
- Poor traceability between test and requirements
- The advantages of requirements for software products are not realized



... with AI-support

Software project development efforts



Idea

AI derives the tests from natural language requirements. This leads to significantly reduced development effort



Setup & Data

Master's Thesis

Partners



VECTOR >



HEICON
Global Engineering



Raw Data Overview

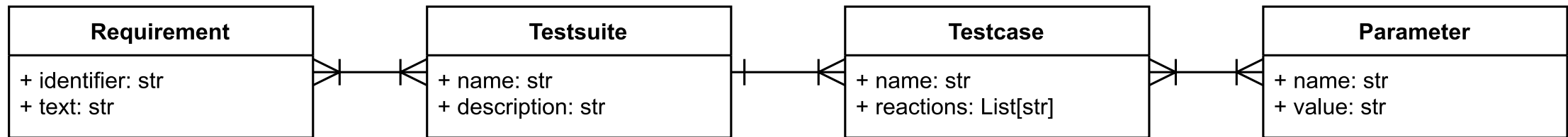
Six projects provided by a renowned aerospace company

Requirements

- 🌀 .docx files
- 🌀 In form of text, tables and images
- 🌀 ~ **6500** requirements

Test specifications

- 🌀 .cte files
- 🌀 Parsable by using a grammar
- 🌀 ~ **4000** testsuites
- 🌀 ~ **9000** unique testcases





Parameter Name Detection

Master's Thesis



Data Labeling

If the toothbrush's battery level drops below 10%, the toothbrush shall activate the energy-saving mode.



Data Labeling

If the toothbrush's **battery level** drops below 10%, the toothbrush shall activate the **energy-saving mode**.



Data Labeling

If the toothbrush's **battery level** drops below 10%, the
toothbrush shall activate the **energy-saving mode**.

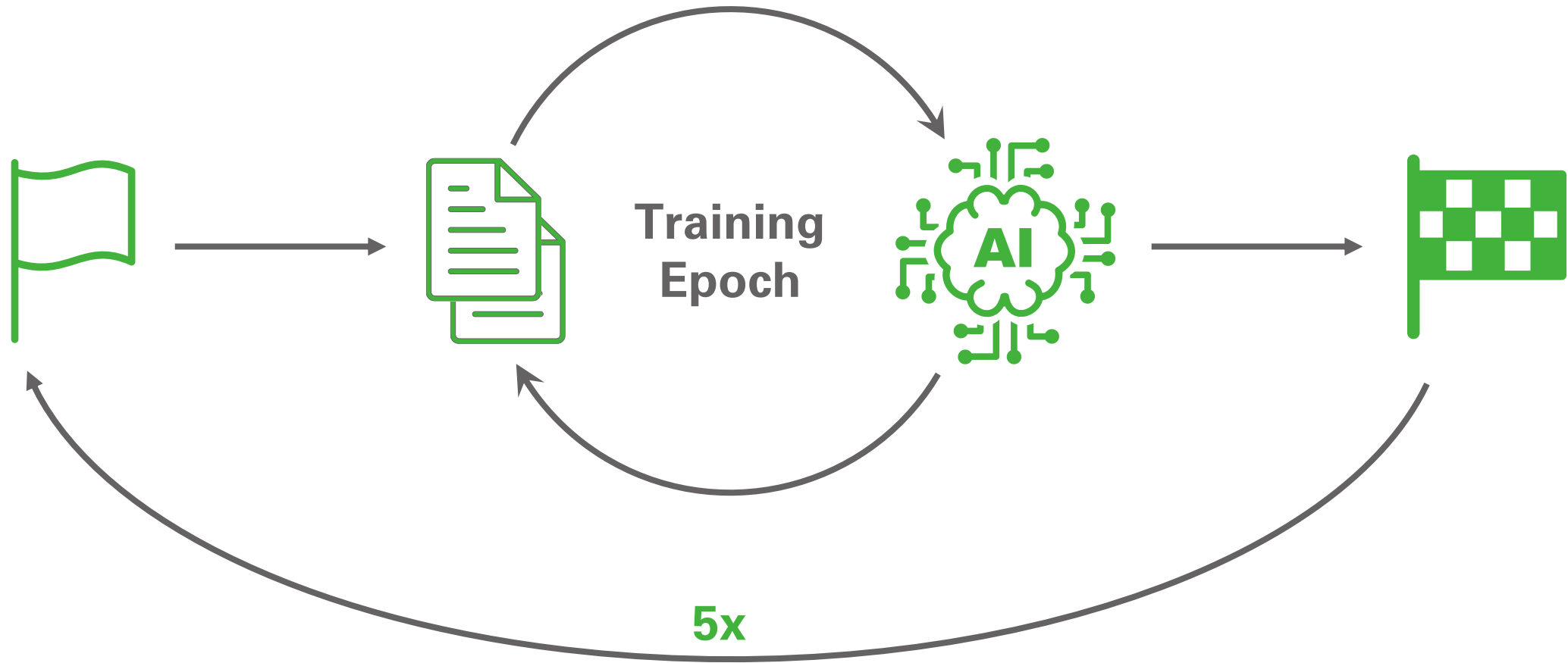


Data Labeling

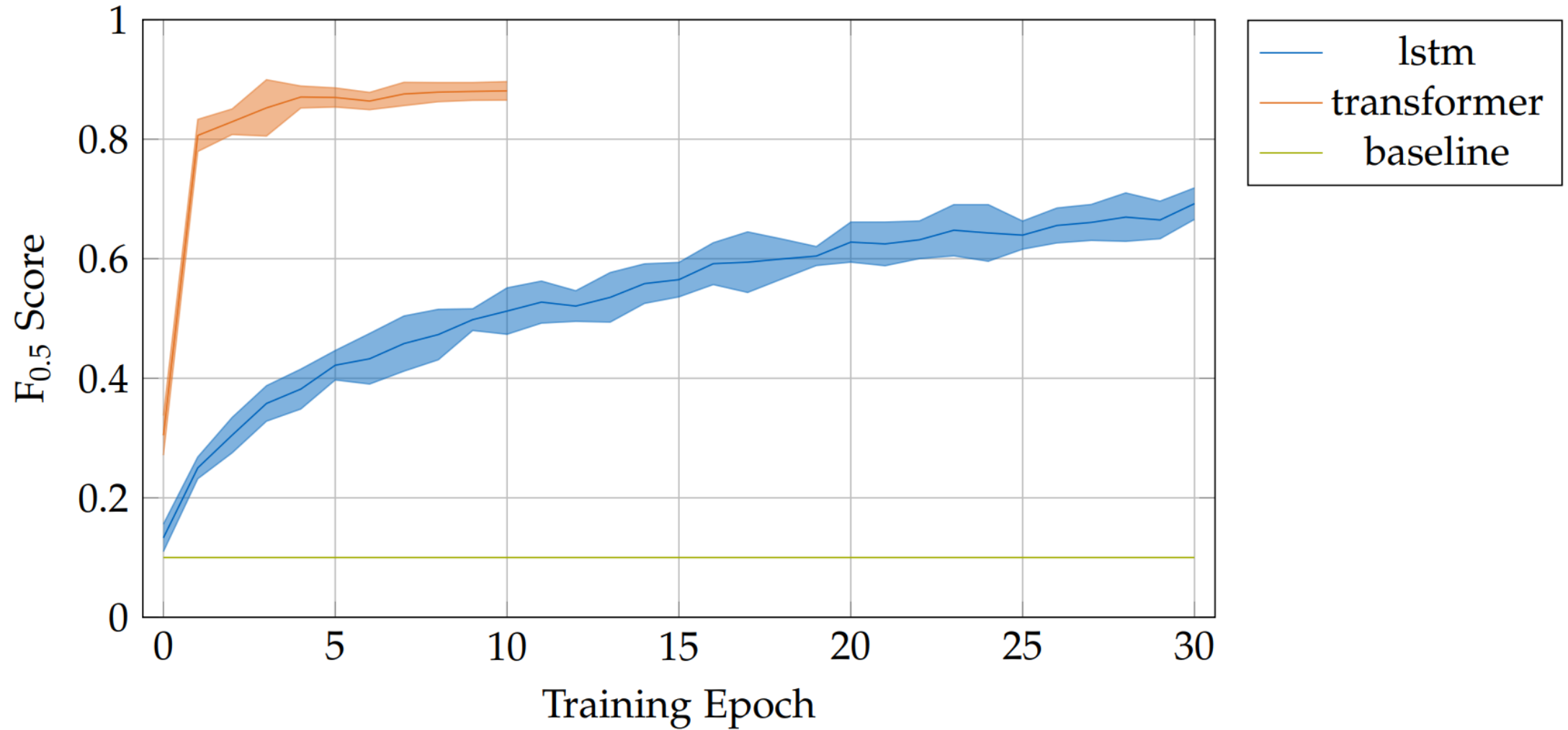
If	the	tooth	brush	's	battery	level	drops	below		10	%,	the
0	0	0	0	0	1	1	0	0	0	0	0	0

tooth	brush	shall	activate	the	energy	-saving	mode	.
0	0	0	0	0	1	1	1	0

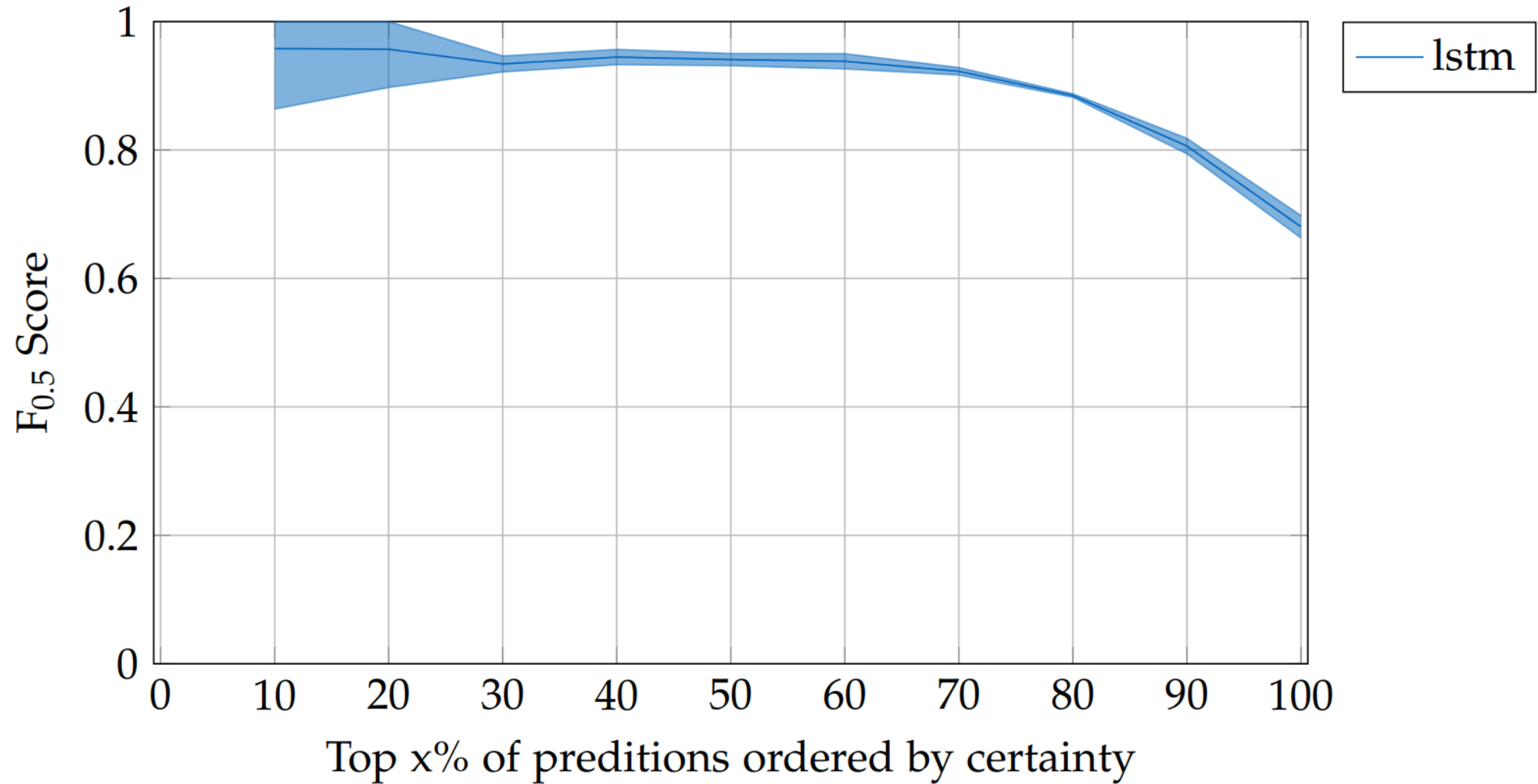
Training Procedure



Networks Performance



Experiment – Certainty





Parameter Value Detection

Master's Thesis



Data Labeling

If the toothbrush's **battery level** drops below 10%, the toothbrush shall activate the **energy-saving mode**.



Data Labeling

If the toothbrush's **battery level** drops below 10%, the toothbrush shall activate the **energy-saving mode**.



Data Labeling

If	the	tooth	brush	's	battery	level	drops	<u>below</u>	_	<u>10</u>	<u>%</u> ,	the
0	0	0	0	0	0	0	0	1	1	1	1	0

tooth	brush	shall	<u>activate</u>	the	energy	-saving	mode	.
0	0	0	1	0	0	0	0	0

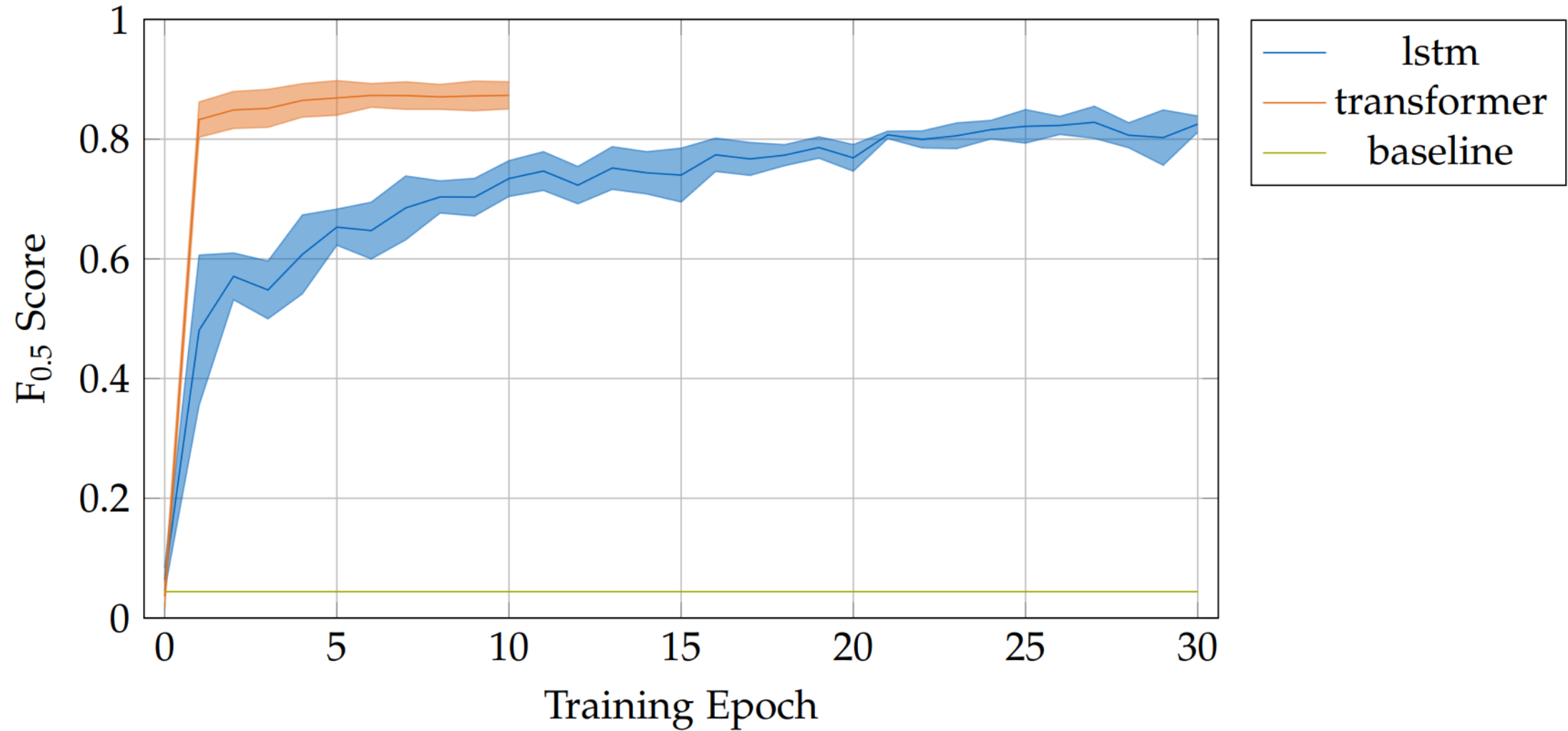


Data Labeling

If	the	tooth	brush	's	battery	level	drops	<u>below</u>	_	<u>10</u>	<u>%</u> ,	the
0	0	0	0	0	1	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	1	1	1	

tooth	brush	shall	<u>activate</u>	the	energy	-saving	mode	.
0	0	0	0	0	1	1	1	0
0	0	0	1	0	0	0	0	0

Networks Performance





Summary

AI supported test generation



Summary

- Master Thesis could demonstrate proof of concept with respect to identify the required data in individual requirements
- A first step in usage the technology can be an analysis tool which highlights weaknesses in requirements
- Different AI algorithms are required to extract the information from natural language, tables and graphs

Topics for further investigations:

- Deriving test scenarios from the data, applying techniques like equivalence class and boundary value testing
- Getting enough high quality requirements to train the AI models to avoid biased results

Statement



“It will be worth continuing to work on this topic”

- Vector Informatik GmbH -



Future Trends

Example: Modern Requirements & Copilot



Modern Requirements: Analysis

5547 If the toothbrush's battery level drops below 10%, the toothbrush shall activate the energy-saving mode.

Analyse

Factor	Rating (1-10)	Reasons for Rating	Suggested Improvement
Clear	8	The requirement is mostly clear but could specify what is meant by "energy-saving mode".	Specify what functionalities are reduced or turned off in energy-saving mode.
Concise	9	The requirement is concise, with no unnecessary words.	Not applicable, as the requirement is already concise.
Complete	6	The requirement does not specify what happens when the battery is recharged above 10%.	Add information about the behavior of the toothbrush once the battery level is back above 10%.
Consistent	10	The requirement uses consistent terminology and is in line with standard requirement formats.	Not applicable, as the requirement is already consistent.
Correct	9	The requirement is technically feasible and relevant to the product's function.	Verify with technical team to ensure 10% is the optimal threshold for activating energy-saving mode.
Concrete	7	The requirement is somewhat vague about what "energy-saving mode" entails.	Define specifically what changes occur in energy-saving mode, such as reduced vibration speed or disabling certain features.

Overall Quality Score Calculation: To calculate the overall quality score, we take the average of the individual ratings: $(8 + 9 + 6 + 10 + 9 + 7) / 6 = 49 / 6 \approx 8.17$ ### Summary of Findings and Recommendations: The requirement for activating the energy-saving mode on the toothbrush when the battery level drops below 10% is generally well-formulated, scoring an overall quality score of approximately 8.17 out of 10. The requirement is clear, concise, consistent, and correct, indicating a strong foundation. However, it lacks completeness and precision, particularly in detailing the behavior of the toothbrush once the battery is recharged above 10% and in specifying what "energy-saving mode" entails. To improve the requirement, it is

Generate

Copy

Done



Modern Requirements: Testcases

5547 If the toothbrush's battery level drops below 10%, the toothbrush shall activate the energy-saving mode.

Given Input: Title, Description Create: Test Case

- Select all
- Verify activation of energy-saving mode when battery level is 9%**
- Test energy-saving mode functionality when battery level is 5%
- Check if energy-saving mode is disabled when battery level is above 10%
- Assess impact of energy-saving mode on toothbrush performance
- Validate battery level monitoring accuracy
- Evaluate energy consumption in energy-saving mode
- Examine battery level threshold for activating energy-saving mode
- Confirm energy-saving mode activation notification to user
- Audit energy-saving mode behavior during charging
- Ensure energy-saving mode does not affect toothbrush functionality

Description
Verify activation of energy-saving mode when battery level is 9%

Test Steps

Step	Action	Expected Results
1	Set the battery level to 9% and trigger the energy-saving mode activation process	The energy-saving mode should be successfully activated when the battery level reaches 9%
2	Monitor the device behavior in energy-saving mode to ensure functionality is optimized	The device should exhibit reduced power consumption and optimized performance while in energy-saving mode

Generate No. of records: Minimal Copy Create linked work item Cancel



Modern Requirements & Copilot



HEICON

Global Engineering



HEICON – Global Engineering GmbH

Member of: BodenseeAIRea, ASQF, Embedded4You

Contact

Martin Heininger
Dipl.-Ing. Electrical Engineering

Phone: +49 7353 981781
Mobile: +49 176 24 73 99 60
martin.heininger@heicon-ulm.de

Contact

August Weiß
Information Systems, M. Sc.

Mobile: +49 151 40 04 49 25
august.weiss@heicon-ulm.de

Offices

In der Neuen Welt 8
87700 Memmingen

Kreuzweg 22
88477 Schwendi



Detailed Results – Parameter Name Detection

	Baseline	LSTM	BERT
F_{0.5}	.100	.692	.881
Accuracy	.818	.943	.935
Precision	.100	.667	.877
Recall	.099	.815	.899
s/epoch	-	3.76	16.8

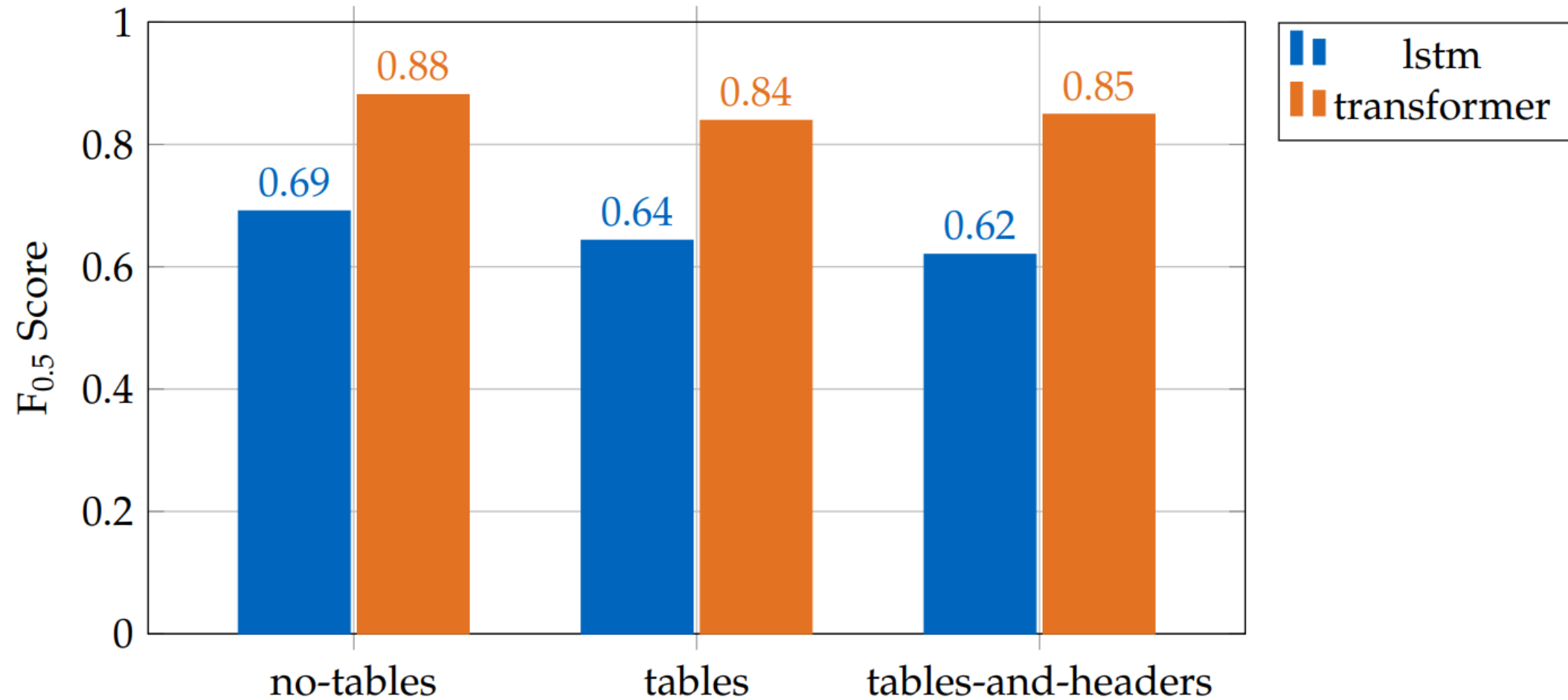


Detailed Results – Parameter Value Detection

	Baseline	LSTM	BERT
F_{0.5}	.044	.825	.873
Accuracy	.917	.988	.993
Precision	.044	.795	.880
Recall	.044	.975	.847
s/epoch	-	11.9	44.3



Experiment - Tables



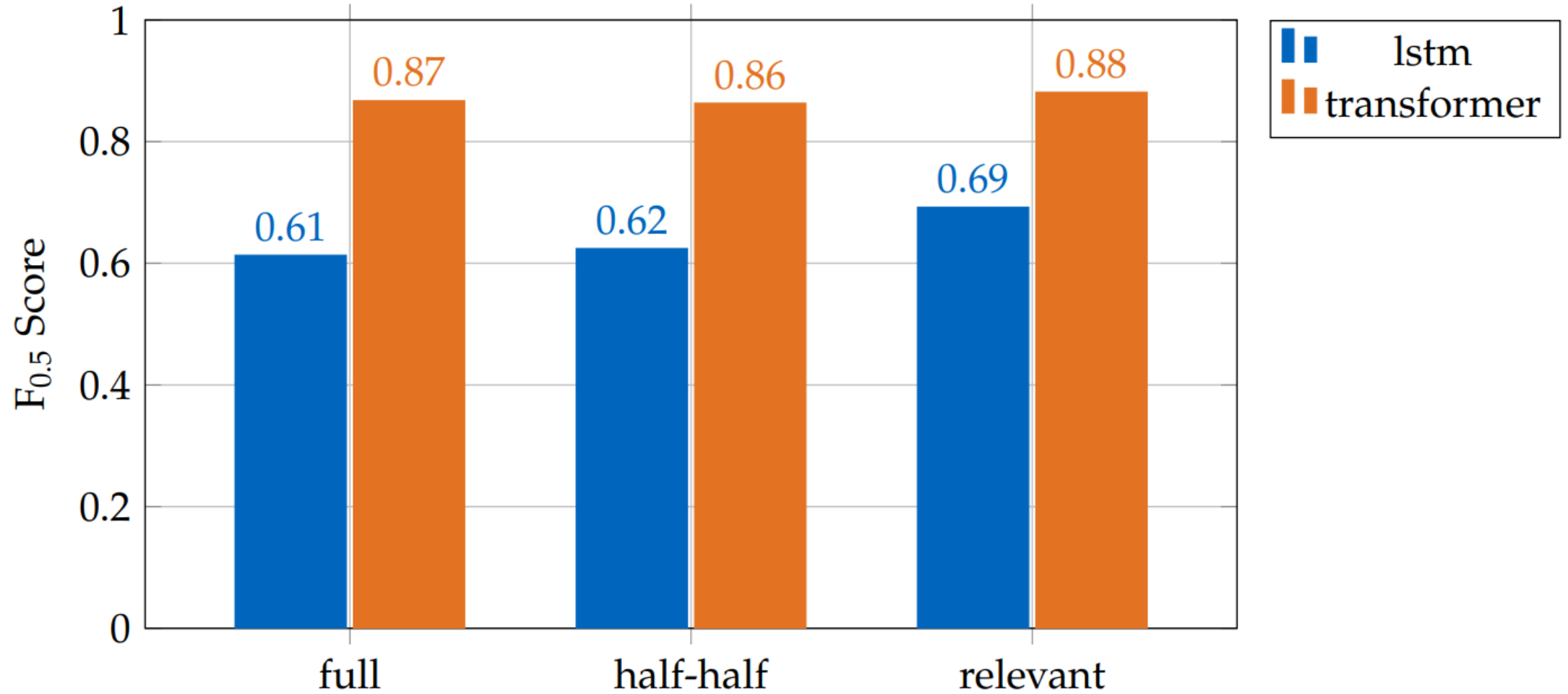


Experiment – Tables

Dataset	Positives	Negatives	Share positives
No tables	7013	62 364	10.1%
Tables	8202	85 850	8.7%
Tables and headers	8422	100 209	7.8%



Experiment – Imbalance





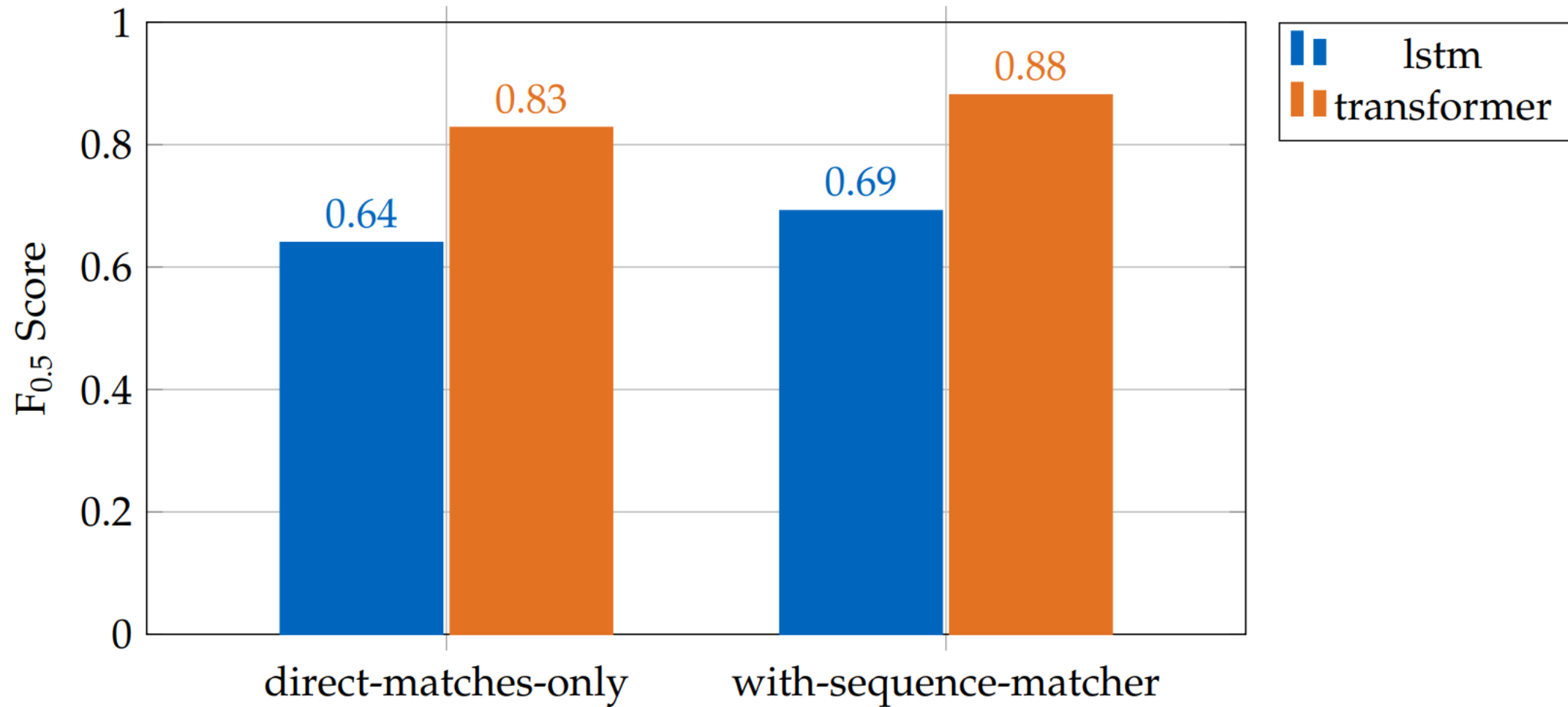
Experiment – Imbalance

Sample-wise	Positives	Negatives	Share positives
Full	1561	1887	45%
Half-half	1561	1561	50%
Relevant	1561	0	100%

Token-wise	Positives	Negatives	Share positives
Full	7013	115 504	5.7%
Half-half	7013	107 174	6.1%
Relevant	7013	62 364	10.1%



Experiment – Sequence Matcher





Experiment – Sequence Matcher

	Direct matches only	With sequence matcher
Positive tokens	4827	7013
Negative tokens	54 151	62 364
Relevant requirements	1286	1561



ChatGPT 3.5



You

Please analyze the following requirement and create a complete test specification for it:

If the industrial oven's temperature exceeds 200°C, the coolant system shall be activated to prevent overheating.



ChatGPT 3.5

1. Introduction:

1.1 Purpose: Verify that the coolant system is activated when the industrial oven's temperature exceeds 200°C to prevent overheating.

1.2 Scope: The testing will cover scenarios where the oven's temperature crosses the 200°C threshold.



ChatGPT 3.5

2. Test Scenarios:

2.1 Oven Temperature Below 200°C:

- ☞ Verify that the coolant system remains inactive when the oven's temperature is below or equal to 200°C.
- ☞ Expected Outcome: Coolant system remains inactive.

2.2 Oven Temperature Exceeds 200°C:

- ☞ Increase the oven temperature beyond 200°C.
- ☞ Expected Outcome: Coolant system should be activated.

2.3 Continuous Monitoring:

- ☞ Continuously monitor the oven temperature as it fluctuates around the 200°C threshold.
- ☞ Expected Outcome: Coolant system activates when the temperature exceeds 200°C.



ChatGPT 3.5

3. Test Cases:

3.1 Oven Temperature Monitoring:

- ☞ Set the oven temperature to 180°C.
- ☞ Monitor temperature readings.
- ☞ Expected Outcome: Coolant system remains inactive.

3.2 Temperature Exceeds 200°C:

- ☞ Set the oven temperature to 210°C.
- ☞ Monitor temperature readings.
- ☞ Expected Outcome: Coolant system activates within a reasonable time.



ChatGPT 3.5

3.3 Gradual Temperature Increase:

- Gradually increase the oven temperature from 180°C to 220°C.
- Monitor temperature readings continuously.
- Expected Outcome: Coolant system activates when the temperature surpasses 200°C.

3.4 Temperature Fluctuations:

- Introduce fluctuations around the 200°C threshold.
- Expected Outcome: Coolant system activates when the temperature exceeds 200°C during fluctuations.



ChatGPT 3.5

4. Environmental Conditions:

4.1 Ambient Temperature:

- ☞ Test the system under varying ambient temperatures.
- ☞ Expected Outcome: The system reacts appropriately under different environmental conditions.

4.2 System Load:

- ☞ Test the system under different load conditions.
- ☞ Expected Outcome: The system remains effective in preventing overheating.



ChatGPT 3.5

5. Performance Testing:

5.1 Response Time:

- Measure the response time from temperature exceeding 200°C to coolant system activation.
- Expected Outcome: Response time within acceptable limits.



ChatGPT 3.5

6. Documentation:

6.1 Test Results:

- Record and document the results of each test case.
- Include details on whether the coolant system activated or remained inactive.

6.2 Issues and Observations:

- Document any issues or unexpected observations during testing.
- Include details on system behavior, error messages, or anomalies.



ChatGPT 3.5

7. Conclusion:

- ☞ Summarize the test results and ensure that the temperature control system meets the specified requirements.