



Conversational Onboarding Mechanism for Project Alignment and Setup Support

Presenters*: Shruti Badrish, Ella Cao, Reese Fairchild, Yashika Kachana

Project Contributors: Anthony Wise, Leilani Battle, Anind Dey

*Denotes Equal Contribution

1 - Motivation

New team members in a research lab often experience a **“knowledge gap”**: a disconnect between information they know and information they need to succeed as a contributing member of the lab. Busy schedules of lab members can make it hard to remedy this and onboard new researchers efficiently.

We aim to develop a **Large Language Model (LLM)** tool to aid and increase the efficiency of this process by capturing and communicating key ideas and claims.

2 - Potential Use Cases



Lab Onboarding



Improving Research Efficiency



Research Collaboration

4 - LLMs for Claim Labeling

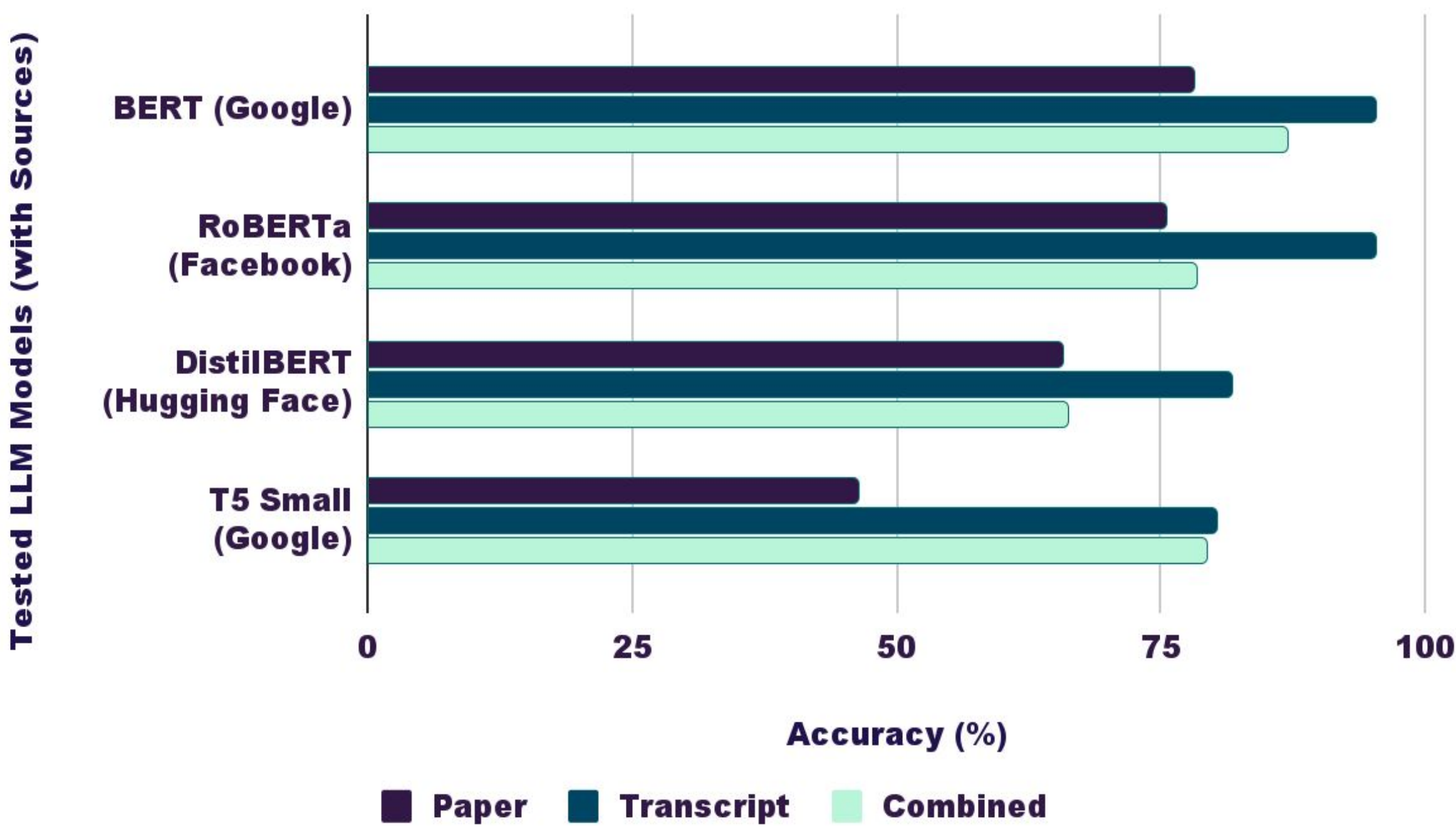
We work to automate the classification of scientific claims with large language models **trained on our classified claims** from Section 3.

We train models on **three** manually classified datasets:

- Claims taken from HCI academic papers (201 claims)
- Claims taken from researcher interview transcripts (328 claims)
- Combined from the other datasets (529 claims)

We measure accuracy by training with **80%** of each dataset and testing with the remaining **20%**.

Accuracy of Trained LLM Models



These results show promise in correctly classifying scientific claims with most accuracies being between **70% - 90%**.

We hypothesize that inaccuracies are caused by our relatively small dataset and because our interview transcripts primarily consisted of “Non-Scientific Claims”; it was not a balanced dataset. Future training should work to mitigate these issues.

3 - Scientific Claims

In order to train an LLM to aid in the research process, we label data from **academic papers** and **interview transcripts** using this codebook:

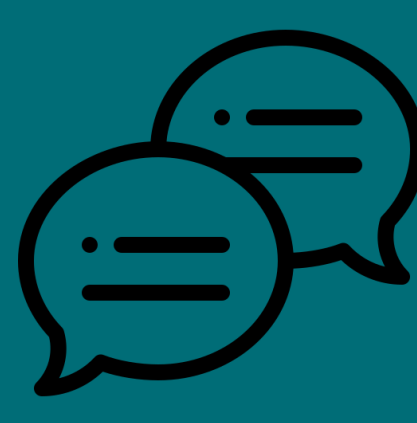
Scientific Claim	Clear, testable assertion. Not a summary of prior work or methods.	e.g. “The Structured format outperformed the Unstructured format, but neither performed impressively.”
Non-Scientific Claim	Not testable, rigorous, or useful claim.	e.g. “Our goal was to see if a system could look at an academic paper and automatically extract the analysis workflow.”
Grey Area Claim	Vague or partial claims.	e.g. “Additionally, there are instances involving improper visual mapping.”
Invalid	Not a syntactically proper sentence.	e.g. “Understood.”

5 - Next Steps

To proceed, we plan to **train and fine-tune a large language model** with our corpus of scientific claim labeling data. Our goal is for this tool to:



Identify key claims and goals from lab material and interviews



Communicate these claims and ideas through a chat feature



Answer clarifying questions that new team members have



Efficiently onboard new research lab members onto projects