RWD44

A large language model (LLM) can **accurately extract genetic testing data** from unstructured clinical records with F-1 scores of 1.0 across multiple fields

Background

- Genetic testing data is often only in unstructured records
- Evidence-generation and clinical care are impeded by inaccessibility of such data
- LLMs perform entity extraction well and show promise for extracting genetic testing data from unstructured records

Aim

Evaluate the performance of an LLM for extraction of genetic test data

Methods

- Gemini 1.5 Pro extracted genetic testing data from genetic counseling notes
- Few-shot prompting with a small number of examples provided in the prompt
- Temperature set to 0 to ensure deterministic outputs
- Label cleaning incorporated into the prompt design for the classification field
- Other hyperparameters, such as top-K and top-N, kept at default values
- Validation process included:

Manual comparison of extracted data against source records

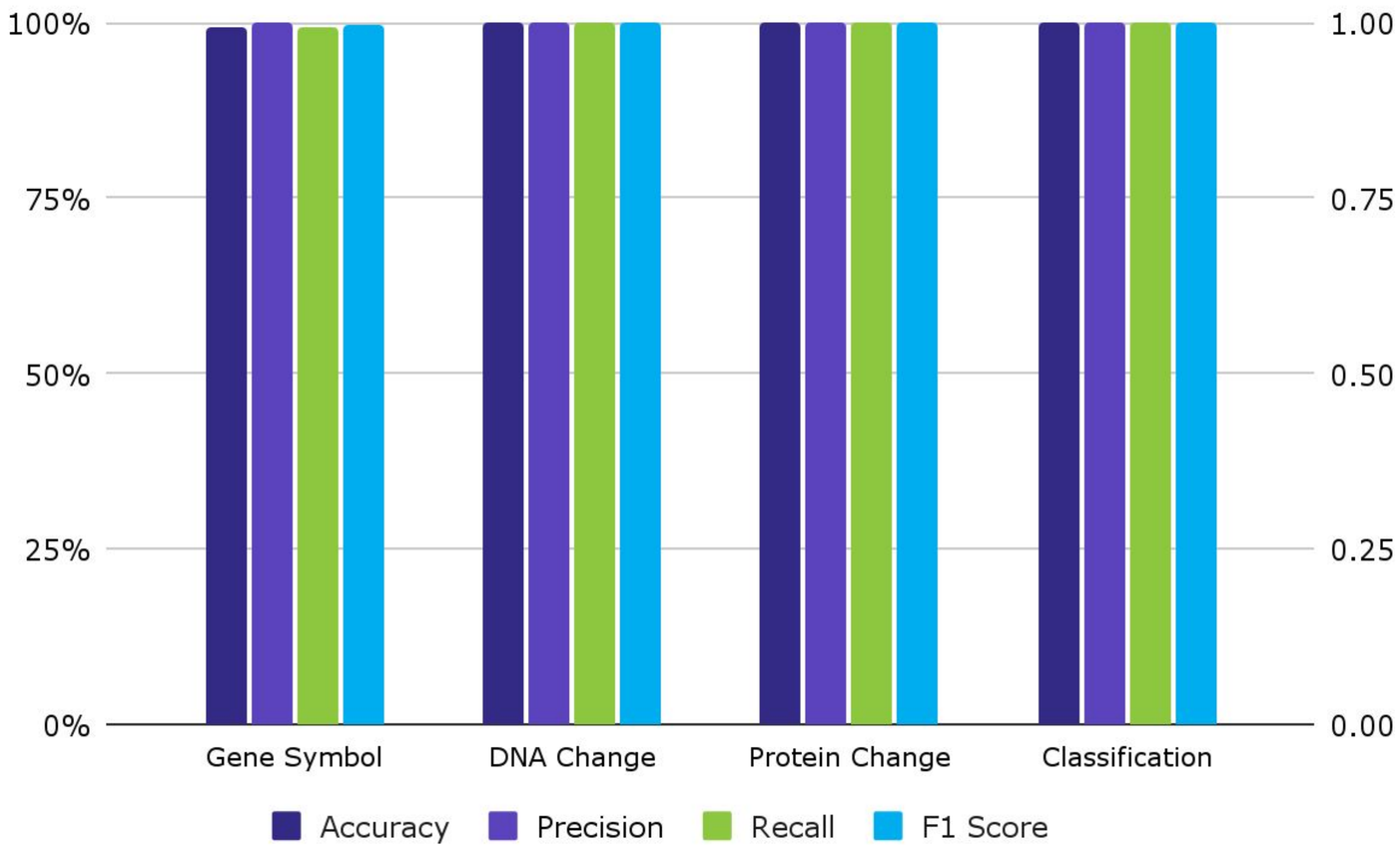
Evaluation of:

Accuracy

Precision

Recall (Sensitivity)

F1 Scores



Results

2692 Genetic Variants

1576 Clinical Records

1557 Patients

96% Of variants were extracted by the LLM

0% Variant hallucination rate by the LLM

Among the variants extracted, the LLM performed well on all metrics across all data fields:

- 99%-100% Accuracy
- 100% Precision
- 99-100% Recall
- 1.00 F1 Scores

Conclusion

- Using an out-of-box LLM to extract genetic data from clinical records is feasible and accurate
- Further work is need to capture the 4% of variants missed by the LLM
- LLMs are a promising approach to improving access to genetic data for healthcare and research purposes

Clinical Validation of Large Language Model for Automated Extraction of Genetic Testing Data

Wenjun He, Colleen Caleshu, Chloe Thorpe, Carra Eagen, Sara Riordan, Andi Hila

Scan for PDF

Contact: colleen@genomemedical.com

authors