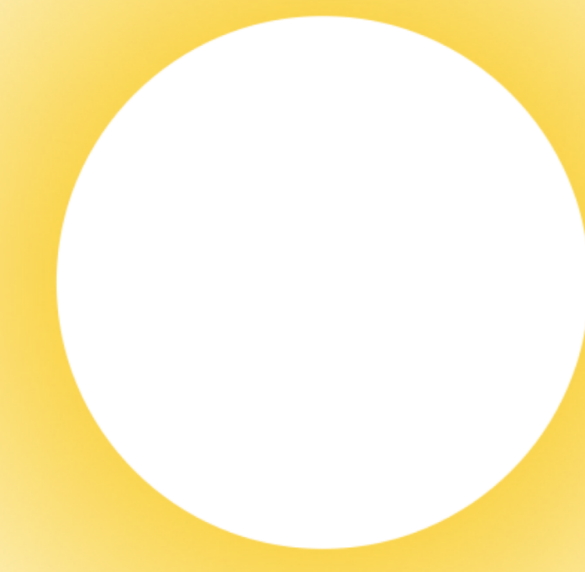


Factiveverse



Vinay Setty
vinay@factiveverse.no
<https://www.factiveverse.ai>

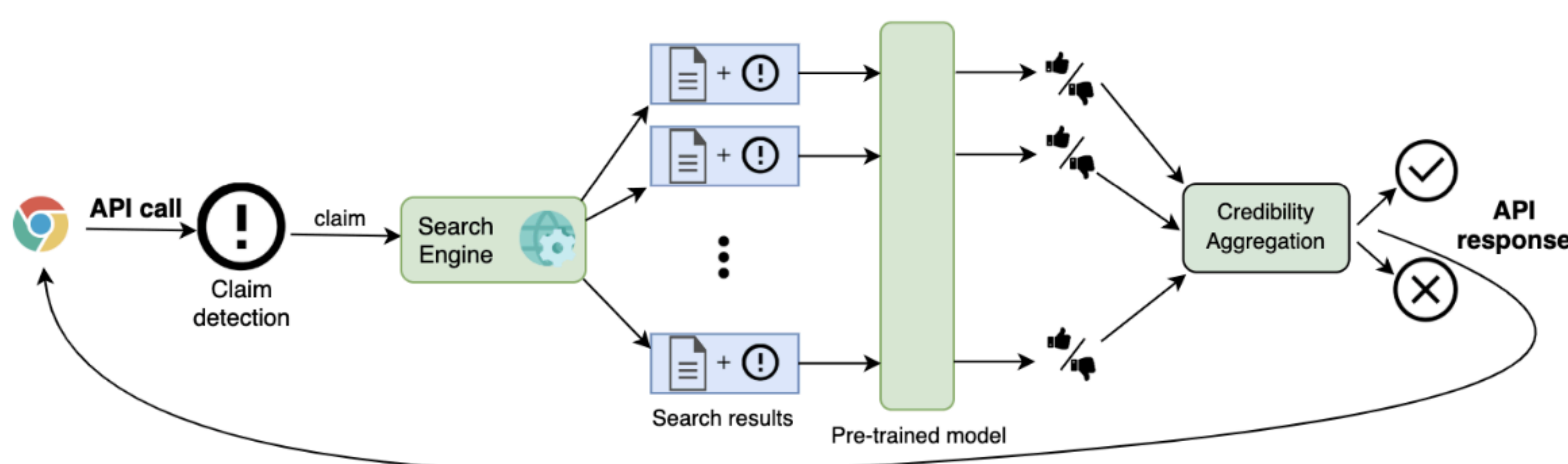
Challenge

Automated fact-checking and research

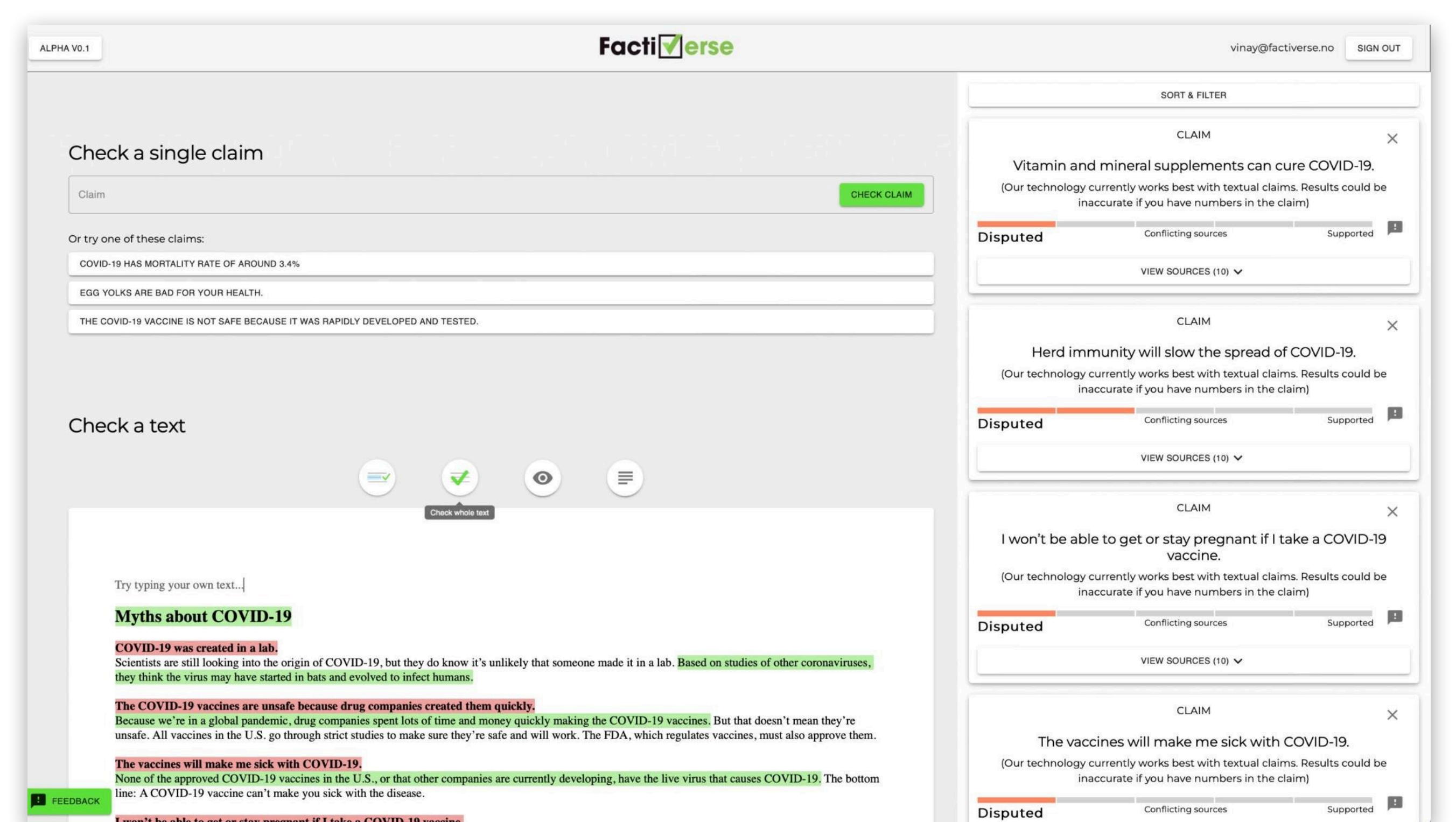
Our Solution



System Architecture



Prototype



Claim Detection

- Identify if a claim is Check-worthy Factual Sentence (CFS), Non Factual Sentence (NFS) or Unimportant Factual Sentence (UFS)
- Features: Word embeddings (W), PoS tags (P), Named entities (N), Dependency tags (D)
- Bi-LSTM and Transformer models
- Trained on US presidential debates

Model	features	f_NFS	f_UFS	f_CFS
CB17 SVM	W_P_N	0.898	0.307	0.696
RF	W_P_N_D	0.804	0.040	0.463
SVM	W	0.833	0.335	0.574
SVM	W_P_N_D	0.836	0.367	0.585
Bi-LSTM	Text_P_N_D	0.867	0.496	0.697
Bi-LSTM	Text_Context	0.890	0.540	0.742
RoBERTa	Text	0.906	0.584	0.769
RoBERTa	Text_Context_P_N_D	0.909 ^Δ	0.574 ^Δ	0.766 ^Δ
Ensemble	All Neural Approaches	0.908	0.589	0.769

Claim Verification

- Trained on manual fact checks (Politifact, Snopes etc.)
- Features: Claim-evidence pairs, metadata (author, subject and domains)
- Bi-LSTM and Transformers

Data	Model	True Acc.	False Acc.	Macro F1	AUC
Politifact	CNN	55.92	57.33	59.39	58.56
	Hi-LSTM	55.85	65.86	60.11	60.66
	Hi-LSTM + Attn.	60.32	68.20	64.80	64.54
	SHAN	62.29	68.43	65.36	65.23
	AHAN	63.25	70.42	66.83	68.66
Snopes	DHAN	60.34	69.76	65.05	65.03
	SADHAN	69.79^Δ	75.45^Δ	71.34^Δ	72.37^Δ
	CNN	72.05	74.29	72.63	76.45
	Hi-LSTM	74.21	74.16	74.33	79.20
Snopes	Hi-LSTM + Attn.	76.76	79.65	77.80	80.33
	DHAN	77.06^Δ	81.63^Δ	78.73^Δ	82.03^Δ

References

- [1] Botnevik, Bjarte, Eirik Sakariassen, and Vinay Setty. "Brenda: Browser extension for fake news detection." In SIGIR 2020
- [2] Mishra, R. and Setty, V., 2019, September. Sadhan: Hierarchical attention networks to learn latent aspect embeddings for fake news detection. In ICTIR 2019.
- [3] Setty V, Rekke E. Truth be Told: Fake News Detection Using User Reactions on Reddit. In CIKM 2019