



DETECTING FACT CHECK- WORTHINESS OF POLITICAL SPEECH

USING BI-DIRECTIONAL RNNs AND TRANSFORMERS

JACOB FRIEDMAN

ITCS 5154 SPRING 2025

LINK TO PRESENTATION: [PRESENTATION ON GITHUB](#)

PROBLEM & CHALLENGES

Problem:

- **Disinformation:** Fact-checking is undeniable in contemporary politics with disinformation abound.
- **Urgency:** Need for rapid, effective verification highlighted (e.g., Trump admin's unsubstantiated claims).
- **Resources:** Issue requires significant resources (news orgs: extensive man-hours/teams for scrutiny).
- **Question (Gencheva et al., 2017):** Can NLP determine fact-worthiness using speech, semantic context, and surrounding text?
- **Project Goal:** Identify claims needing fact-checking via text and semantic analysis.

Challenges:

- Extensive hyperparameter tuning required.
- Limited computational resources (local machine & Google Colab).
- Integrating complex preprocessed data from original authors with new architecture.
- Addressing inherent data imbalance.

MY MOTIVATION

Core Motivation: Driven by palpable political uncertainty and the pervasive influx of misinformation/fabricated narratives by politicians.

The Need: Critical demand for accessible, efficient ways to identify claims needing fact-checking amidst overwhelming volume.

Interdisciplinary Approach: A multi-disciplinary perspective, combining my Political Science background with NLP techniques.



EXISTING RELATED APPROACHES

- **Gencheva et al. (2017): Context-Aware Claim Detection**

- Used State Vector Machine (SVM) and Feed-Forward Networks (FNN) on annotated political debate data
- Created dataset CW-USPD-2016 “Check-Worthiness in the US Presidential Debates 2016”
- Leveraged rich context (debate, speaker interactions, reactions).
- Demonstrated superior performance to baseline, highlighting context importance.
- Limitations: Data dependency, complex feature engineering, potential annotation bias.

- **Other Approaches (Indirectly Relevant):**

- **Chen et al. (2021) (Social Bots):**

- Used neutral "drifter" bots on Twitter to study political bias, concluding bias stemmed from user interactions/platform mechanisms, not the algorithm itself.

- **Ash et al. (2024) (Narrative Analysis):**

- Introduced RELATIO (using SRL + entity clustering) to quantify narrative structures; applied it to the U.S. Congressional Record to analyze dynamics, sentiment, & polarization over time.



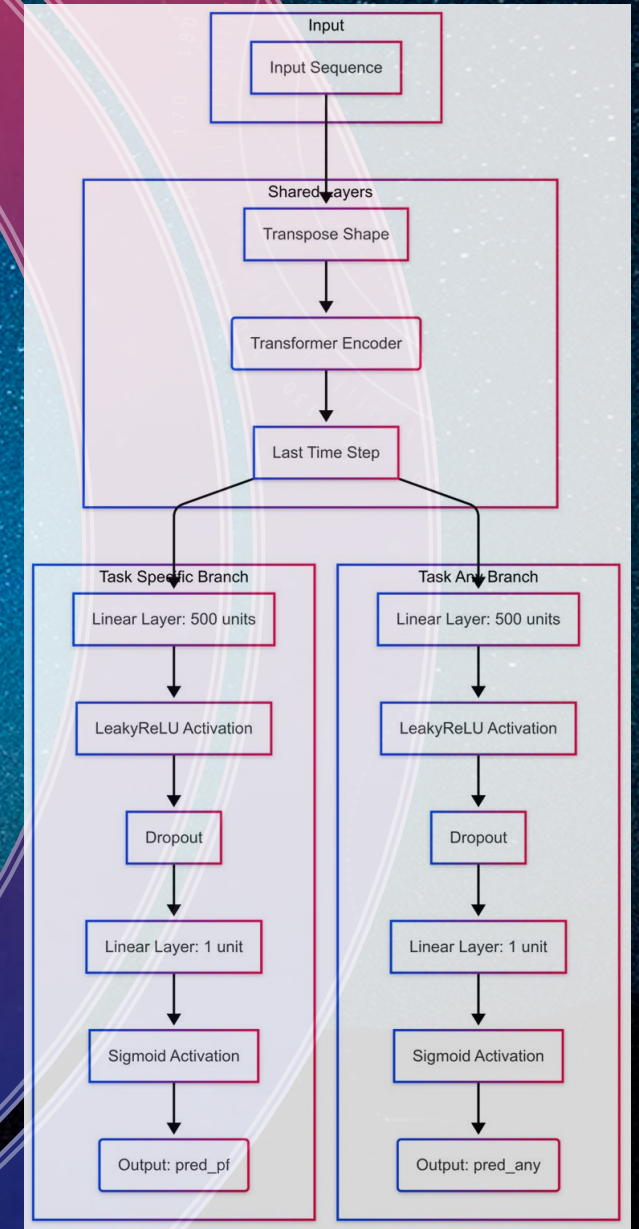
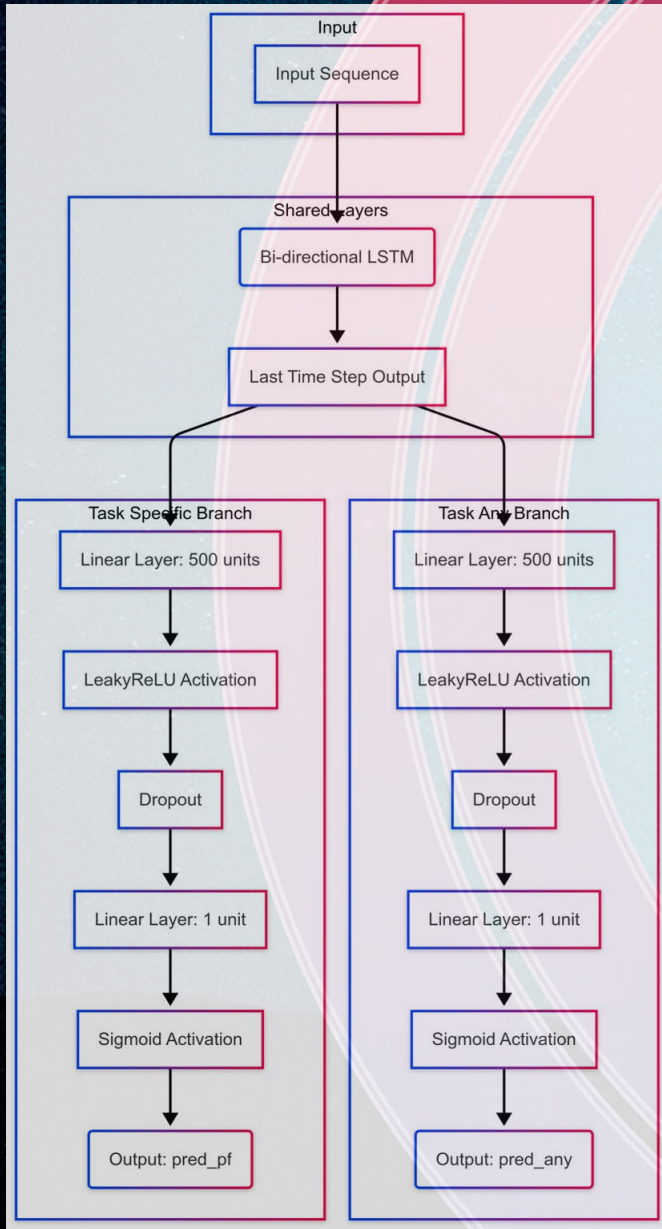
MY METHOD

- **Goal:** Reimplement Gencheva et al. (2017) using Recurrent Neural Network (RNN) & Transformer (vs. original SVM/FNN design).
- **Dataset:** CW-USPD-2016:
 - 4 2016 US election debates (3 presidential, 1 VP)
 - 5,415 sentences annotated by 9 fact-checkers (e.g., CNN, PolitiFact).
 - Low inter-source agreement highlights ranking task focus.
- **Comparability:** Used same dataset & replicated original preprocessing for fair comparison.
- **Key Change:** Replaced the initial SVM layer with:
 - LSTM layer (RNN model: **MultiTaskRNN**).
 - Transformer encoder layer (Transformer model: **MultiTaskTransformer**).
- **Architecture:** Maintained original subsequent layers (Linear, Leaky ReLU, Sigmoid) & task branching.

ARCHITECTURE DIAGRAMS

← MultiTaskRNN Model

MultiTaskTransformer Model →



EXPERIMENTAL STEPS



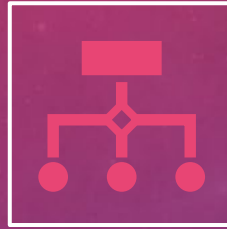
1. Data Pre-processing:

Preprocessing Pipeline: Replicated original authors' complex pipeline using their code

Data Loading: Loaded raw debate transcripts; partitioned into train/validation sets.

Feature Extraction: Converted text to numerical format (using context, POS tags, style features).

Label Extraction: Extracted check-worthiness labels from various sources (e.g., NPR, CNN, PolitiFact).



2. Model Implementation:

Implemented two architectures based on Gencheva et al (2017):

- **MultiTaskRNN:** Shared LSTM layer for temporal dependencies, followed by task-specific branches (Linear, Leaky ReLU, Dropout, Sigmoid).
- **MultiTaskTransformer:** Shared Transformer Encoder using self-attention, followed by similar task-specific branches.



3. Model Training & Evaluation:

Hyperparameter Tuning: Systematic Bayesian Optimization (w/ k-fold CV & early stopping).

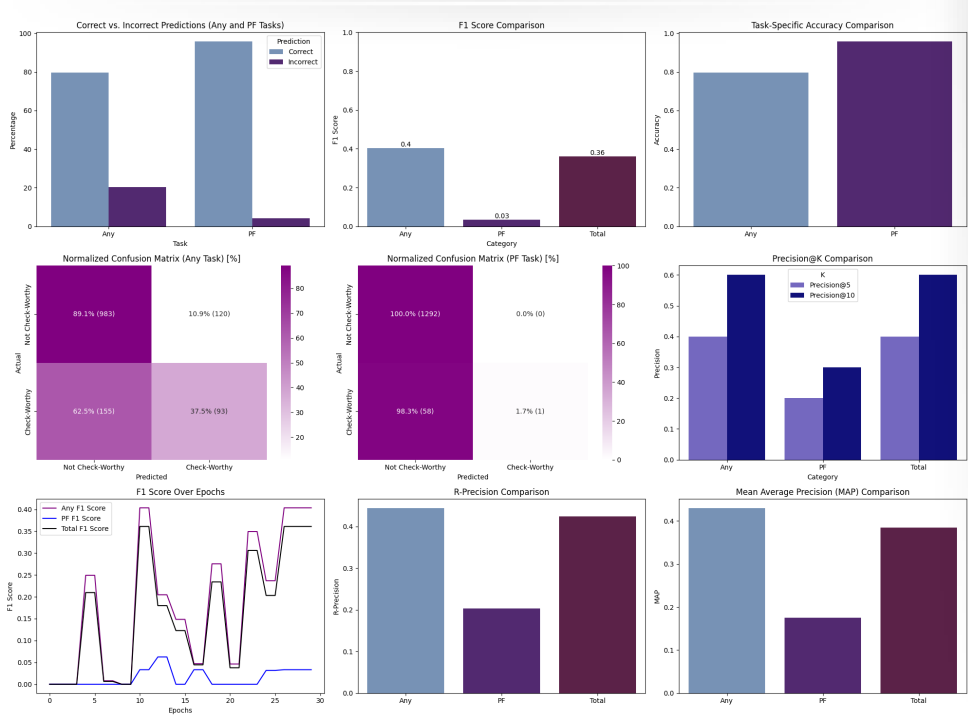
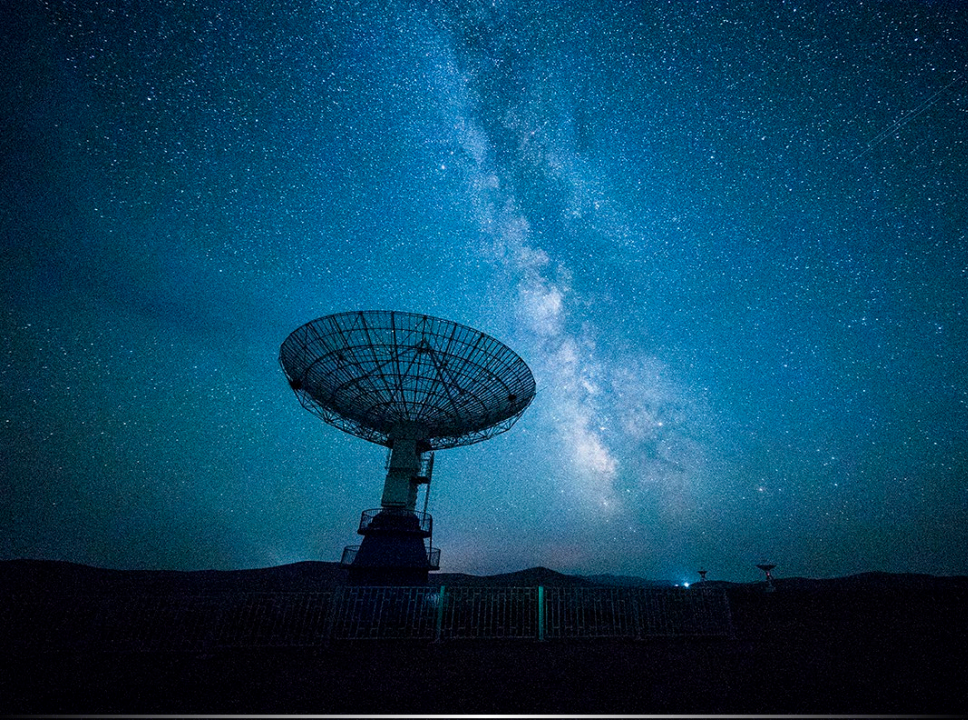
Tuning Phases (Iterative): Broad search -> Optimizer focus -> Architecture focus

Training: Used optimal hyperparameters & Binary Cross-Entropy loss (with multi-task variations).

Evaluation: Assessed final models on held-out validation set (analyzing metrics & visualizations).

RESULTS & OBSERVATIONS

- **Best Performing Model:**
 - Un-Weighted MultiTaskTransformer w/ Combined Score validation.
 - Highest F1 scores for "Any" (~0.40) & "Total" (~0.36) tasks -> better precision/recall.
 - Improved "Check-Worthy" identification for "Any" task (37.5% correct).
 - Enhanced Precision @ K (more relevant claims in top predictions).
- **General Observations:**
 - Accuracy often misleading due to class imbalance; F1 score more reliable.
 - All models struggled significantly with the "PF" (PolitiFact) task.
 - Combined score validation generally boosted F1 performance over average loss (esp. for Transformer).
 - Task weighting showed inconsistent benefits; unweighted was best for top Transformer model.
 - Ranking performance (MAP scores) didn't improve as much as F1 scores.



CONCLUSION & FUTURE WORK

- **Conclusion:**

- Best model (Un-Weighted Transformer w/ Combined Score) is promising, but challenges persist.
- Improved precision/recall achieved for general ("Any") claim detection.
- Needs Improvement: Performance on "PF" task and overall claim ranking.

- **Future Work:**

- Explore advanced evaluation metrics (beyond average loss).
- Investigate dynamic/adaptive task weighting.
- Apply specialized loss functions (e.g., focal loss) for class imbalance (esp. "PF" task).
- Explore RNN/Transformer architecture variations (e.g., hybrids) & further optimization.



Works Cited:

- Ash, E., Gauthier, G., & Widmer, P. (2024). Relatio: Text Semantics Capture Political and Economic Narratives. Political Analysis , 115-132.
doi:<https://doi.org/10.1017/pan.2023.8>
- Chen, W., Pacheco, D., Yang, K.-C., & Menczer, F. (2021). Neutral Bots Probe Political Bias on Social Media. Nature Communications, 12 , 5580.
doi:<https://doi.org/10.1038/s41467-021-25738-6>
- Gencheva, P., Koychev, I., Márquez, L., Barrón-Cedeño, A., & Nakov, P. (2017). A Context-Aware Approach for Detecting Check-Worthy Claims in Political Debates. Proceedings of Recent Advances in Natural Language Processing (pp. 267–276). Varna: INCOMA Ltd.
doi:https://doi.org/10.26615/978-954-452-049-6_037

THANK YOU!

For more details and the full experimental accounting:

Check out my GitHub here:

[ITCS5154CourseProject on GitHub](#)

Check out my full report here:

[PDF of Full Report on GitHub](#)

Check out my presentation here:

[Presentation Video on GitHub](#)