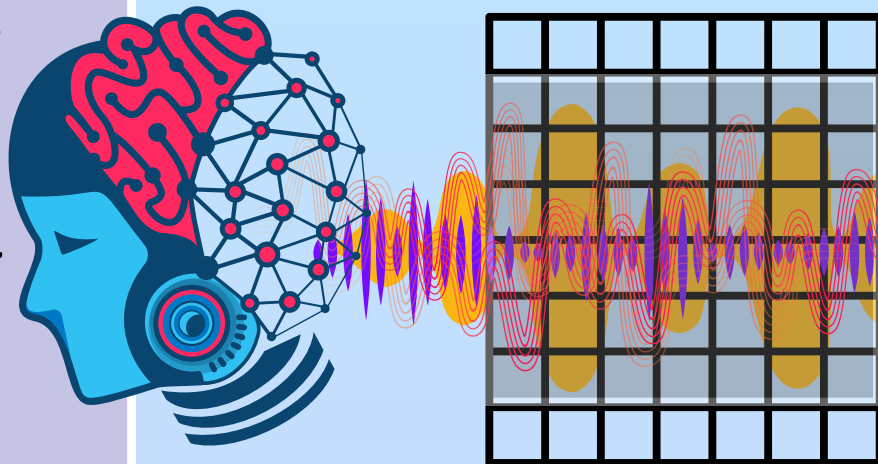


6 Conclusion & Future Research Trajectories

Future Trajectories:

- **Multi-Modal Integration:** Fusing voice with facial and text data.
- **Continuous Emotion Mapping:** Moving beyond discrete labels to valence-arousal space.
- **Real-Time Deployment:** Developing low-latency models for live applications.

Our work redefines speech emotion recognition as a dynamic learning problem, achieving a new standard of adaptability in human-centered AI.



Plot

1 Beyond Static: The Future of Emotion Recognition is Adaptive

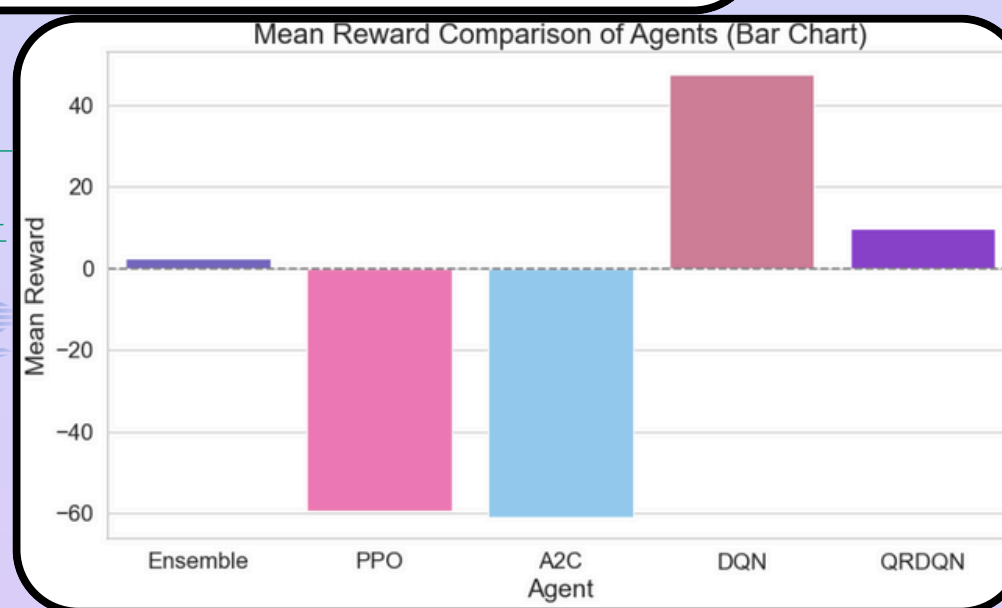
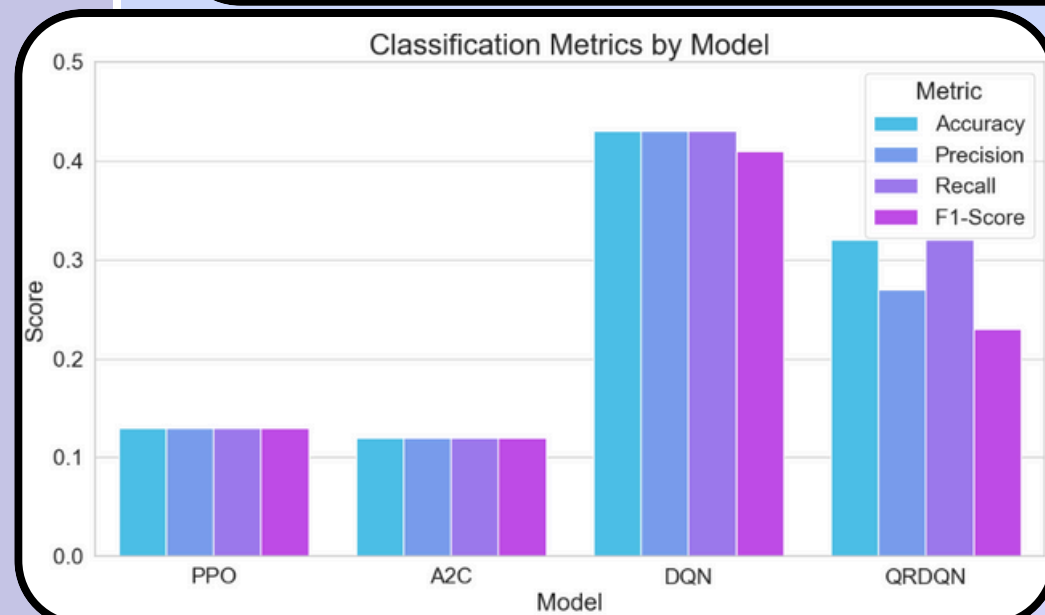
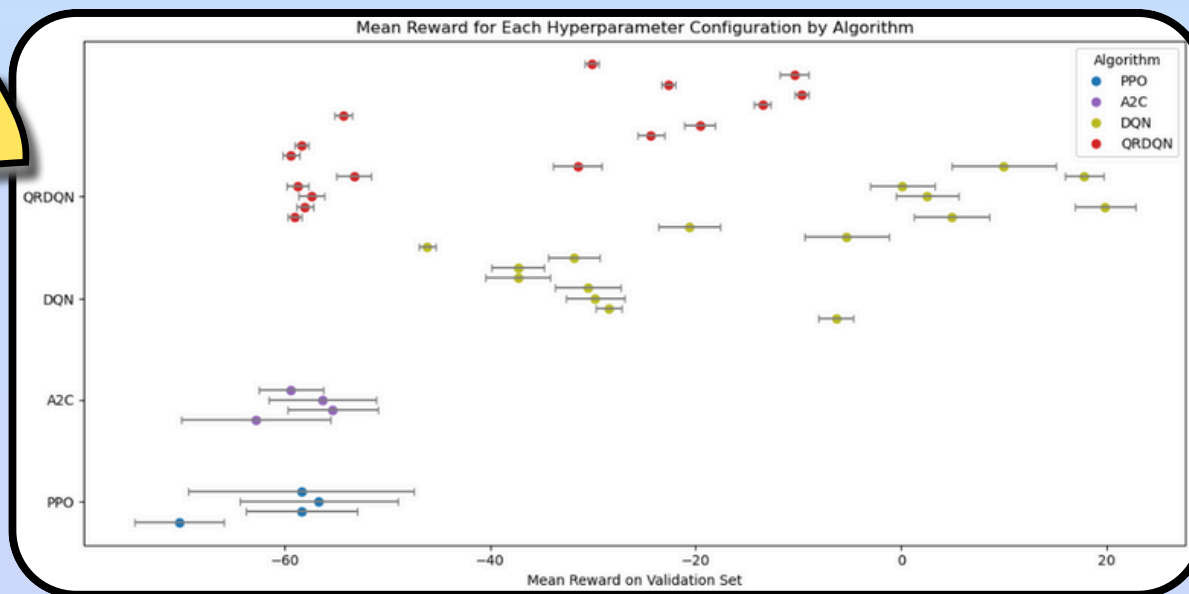
Old Way: Static, rigid models that break.

Our Way: Dynamic, adaptive learning that listens.

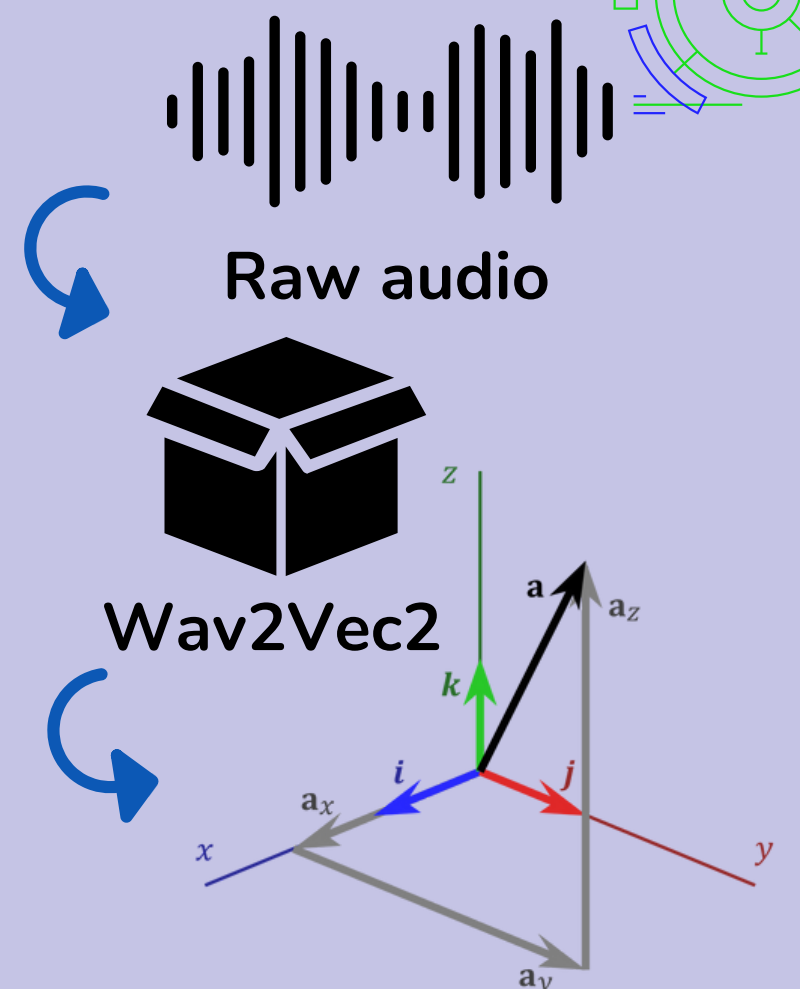
5 Explore Our Results

Take control of the data. Use the central selector to dive into our empirical validation:

- **Analyse:** Hyperparameter Sensitivity
- **Benchmark:** Classification Performance
- **Track:** Agent Learning Dynamics



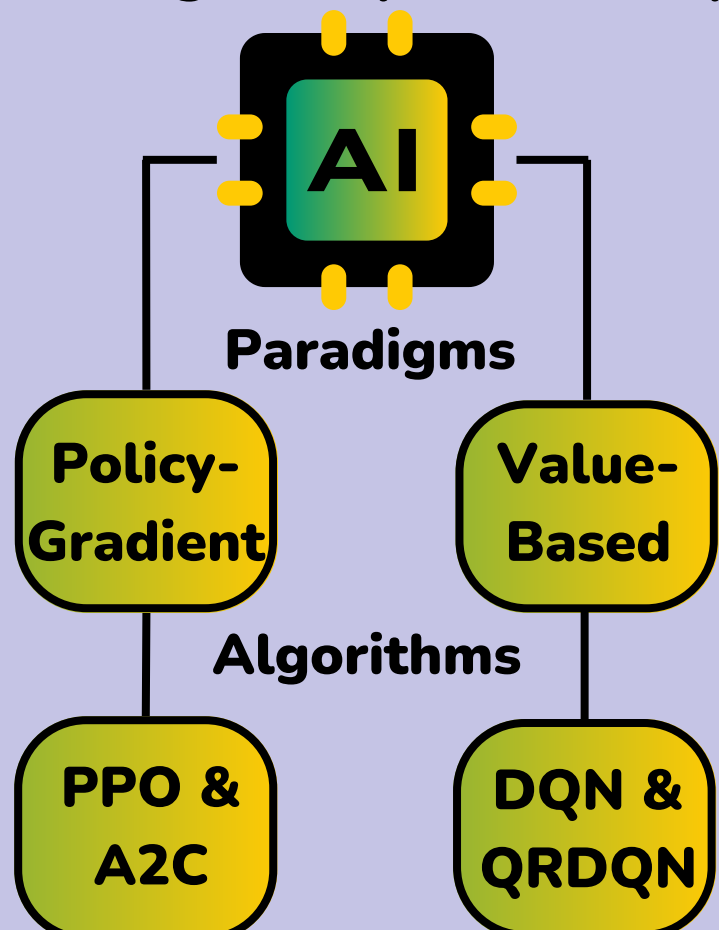
2 Unlocking High-Fidelity Sound with Self-Supervised AI



Contextualized Embedding

4 Advanced Policy Optimization

Crafting the Optimal Policy



3 Formulating SER as a Markov Decision Process (MDP)

- **State (S):** The sequence of acoustic embeddings derived from the current speech segment.
- **Action (A):** The agent's selection of a discrete emotion label from the predefined emotion set.
- **Reward (R):** A correct prediction gives a +1.0 base reward, with an additional +0.2 × confidence bonus. An incorrect prediction results in a -0.5 penalty.

