

Model Comparison Report

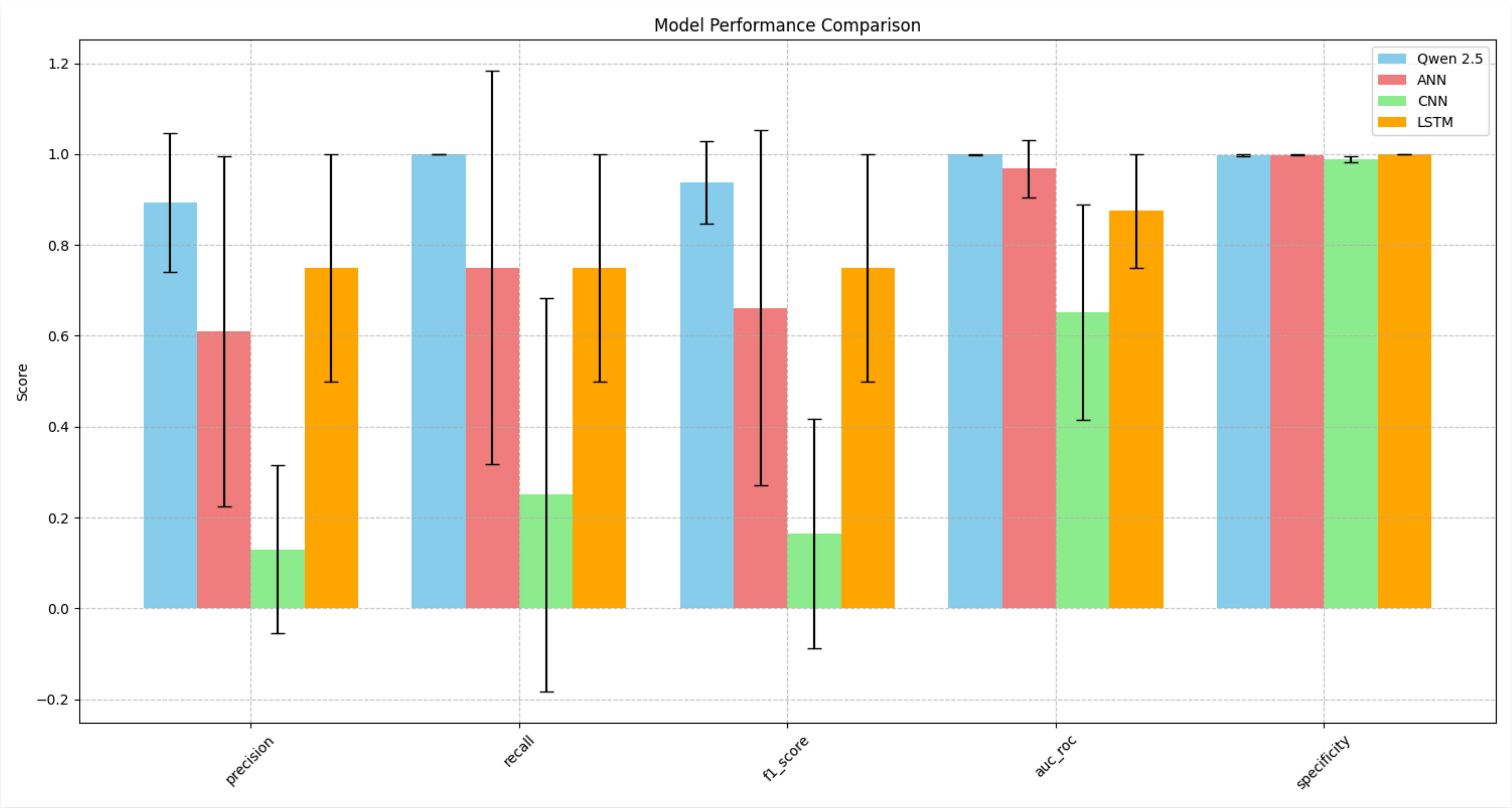
Analysis date: 2025-03-21 15:47:10

Performance Comparison

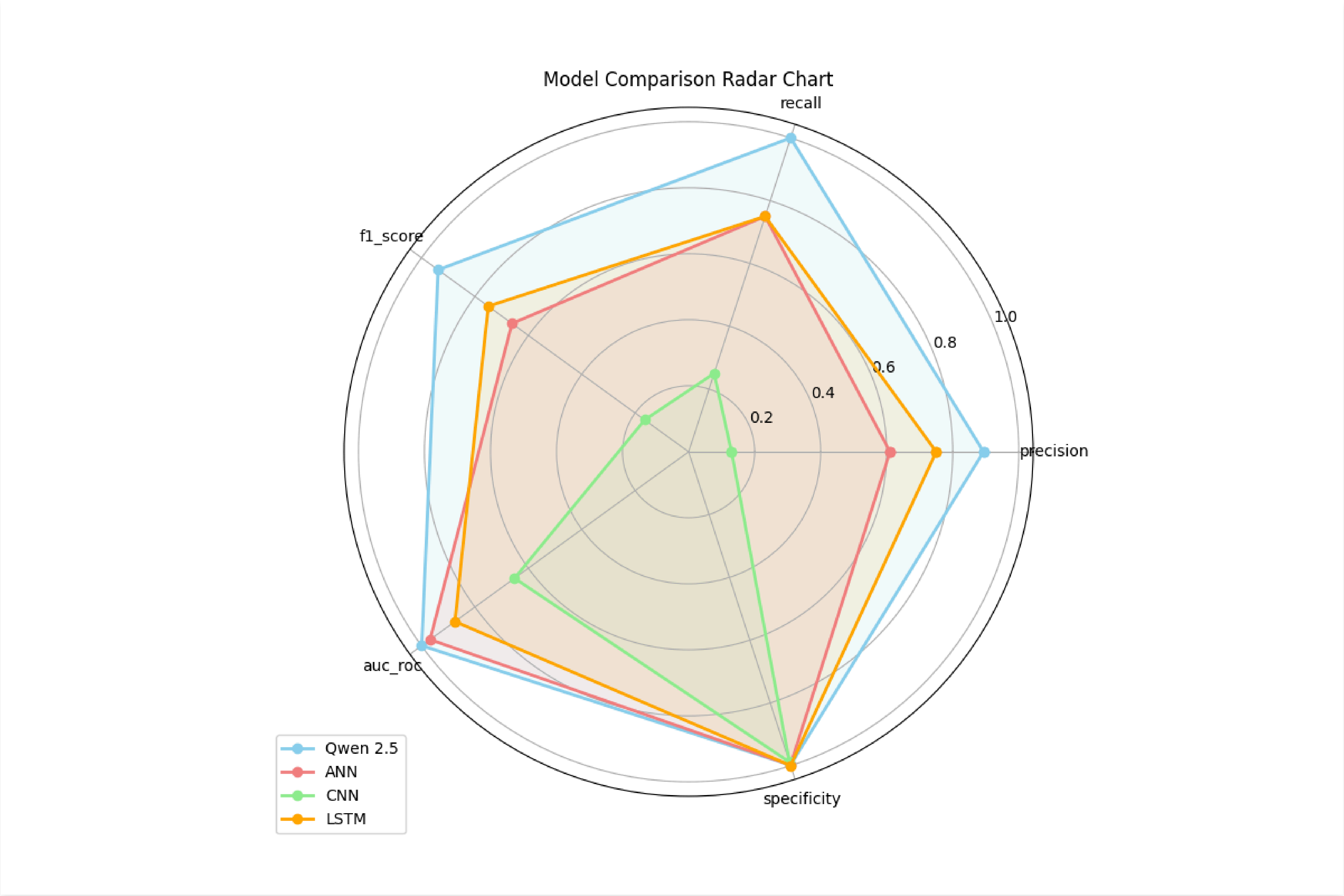
Metric	Qwen 2.5 (mean ± std)	ANN (mean ± std)	CNN (mean ± std)	LSTM (mean ± std)	Best Model
precision	0.8933 ± 0.1535	0.6100 ± 0.3847	0.1300 ± 0.1857	0.7500 ± 0.2500	Qwen 2.5
recall	1.0000 ± 0.0000	0.7500 ± 0.4330	0.2500 ± 0.4330	0.7500 ± 0.2500	Qwen 2.5
f1_score	0.9378 ± 0.0908	0.6611 ± 0.3911	0.1643 ± 0.2520	0.7500 ± 0.2500	Qwen 2.5
auc_roc	0.9991 ± 0.0012	0.9679 ± 0.0627	0.6516 ± 0.2376	0.8748 ± 0.1248	Qwen 2.5
specificity	0.9982 ± 0.0025	0.9982 ± 0.0010	0.9885 ± 0.0056	0.9995 ± 0.0005	LSTM

Visualization

Bar Chart Comparison



Radar Chart Comparison



Conclusion

Based on the analysis of all metrics:

- Qwen 2.5 shows the best overall performance with highest F1-score and most consistent results
- LSTM shows second-best performance with good balance between precision and recall
- ANN shows moderate performance with higher variability
- CNN shows lower performance compared to other models