Model Performance Comparison: Normal vs. Worst-Case

