

TracIn Influence Analysis Report

Plastic 500Hz Dataset

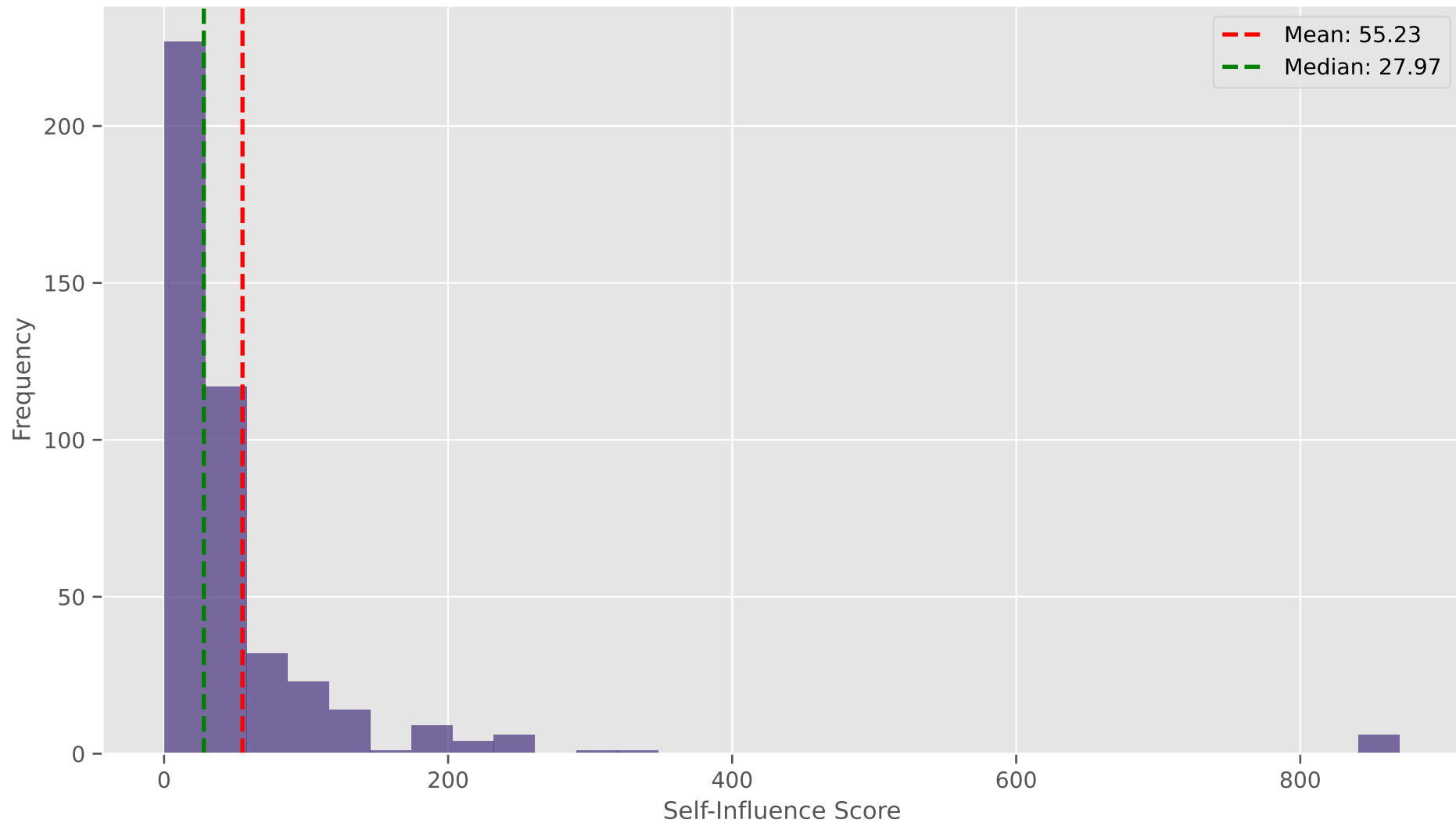
Generated on 2025-04-20 23:25:58

Based on "Estimating Training Data Influence
by Tracing Gradient Descent"

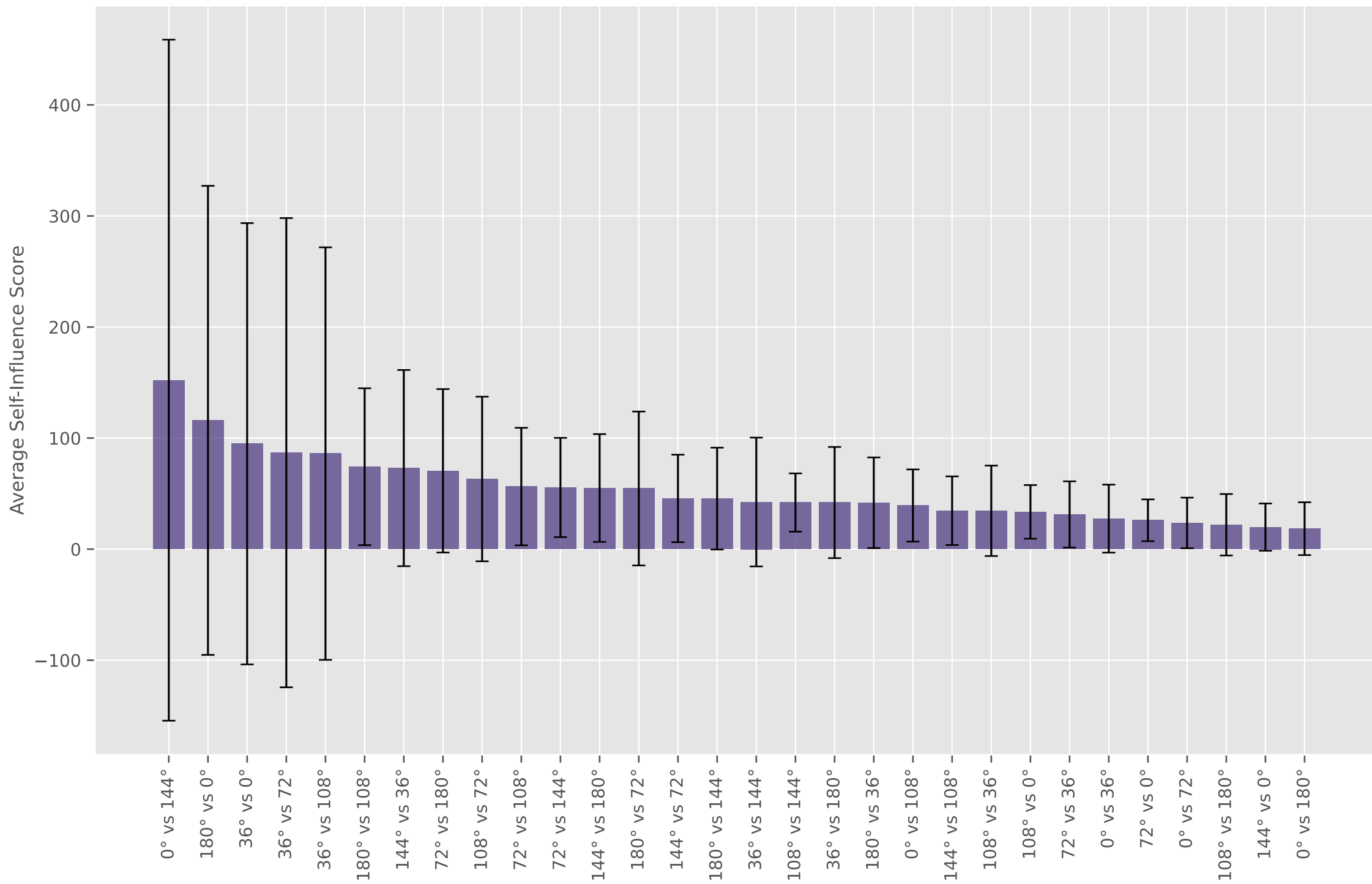
Self-Influence Analysis

Total Samples: 441

Distribution of Self-Influence Scores



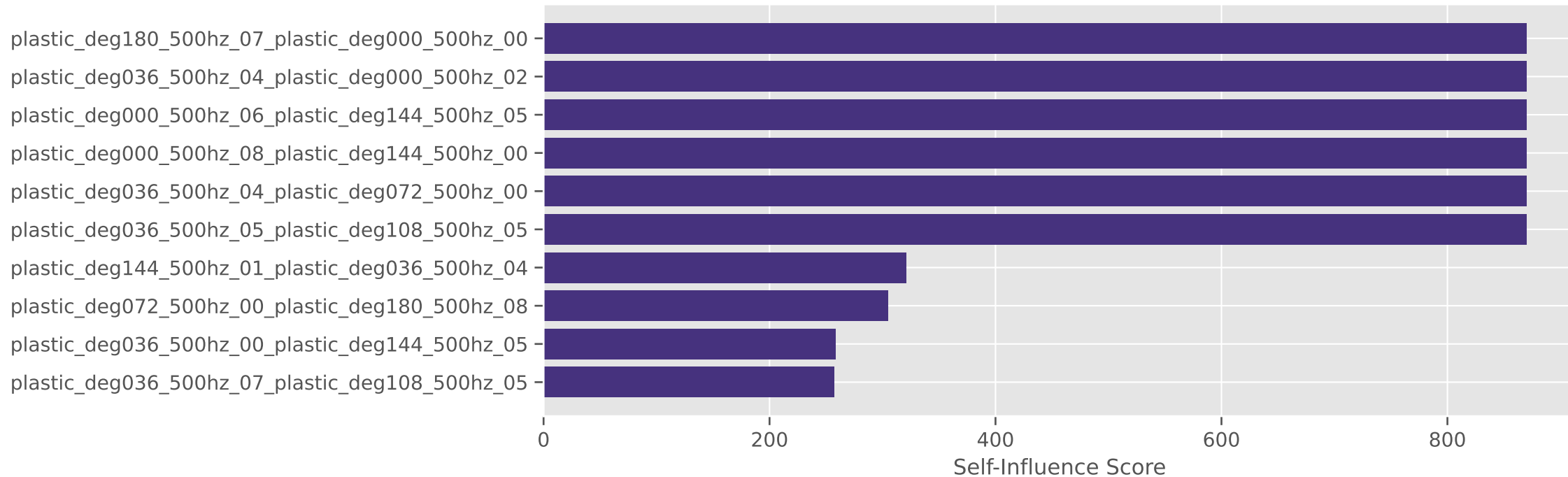
Self-Influence Scores by Degree Pair



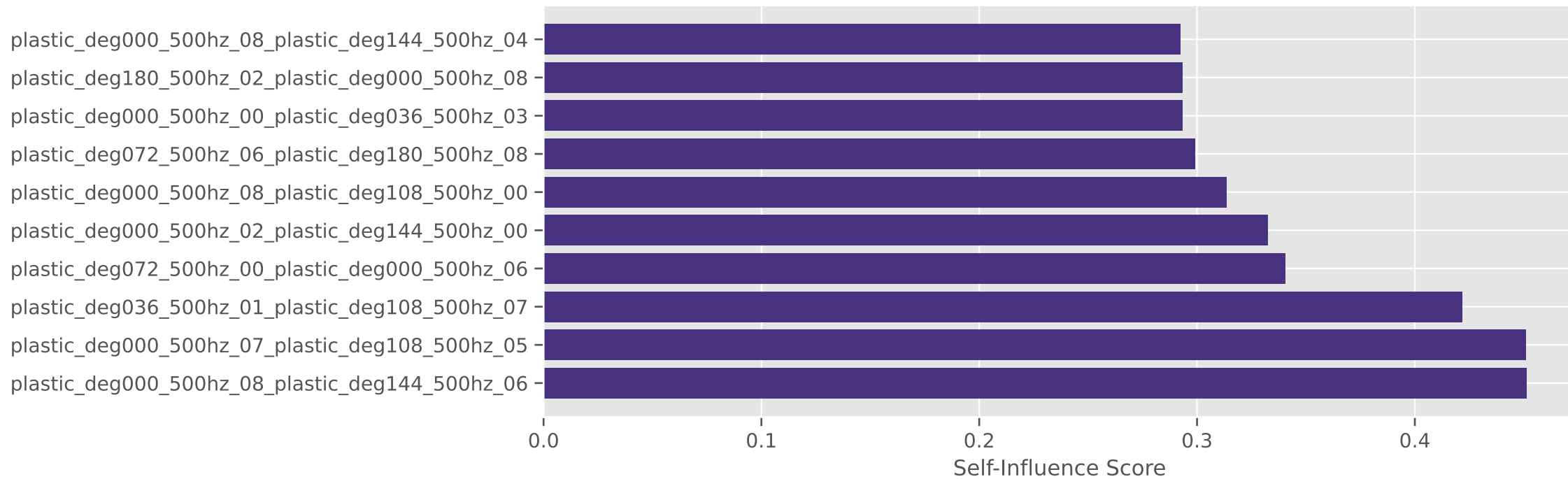
Average Self-Influence Score by Degree Combination



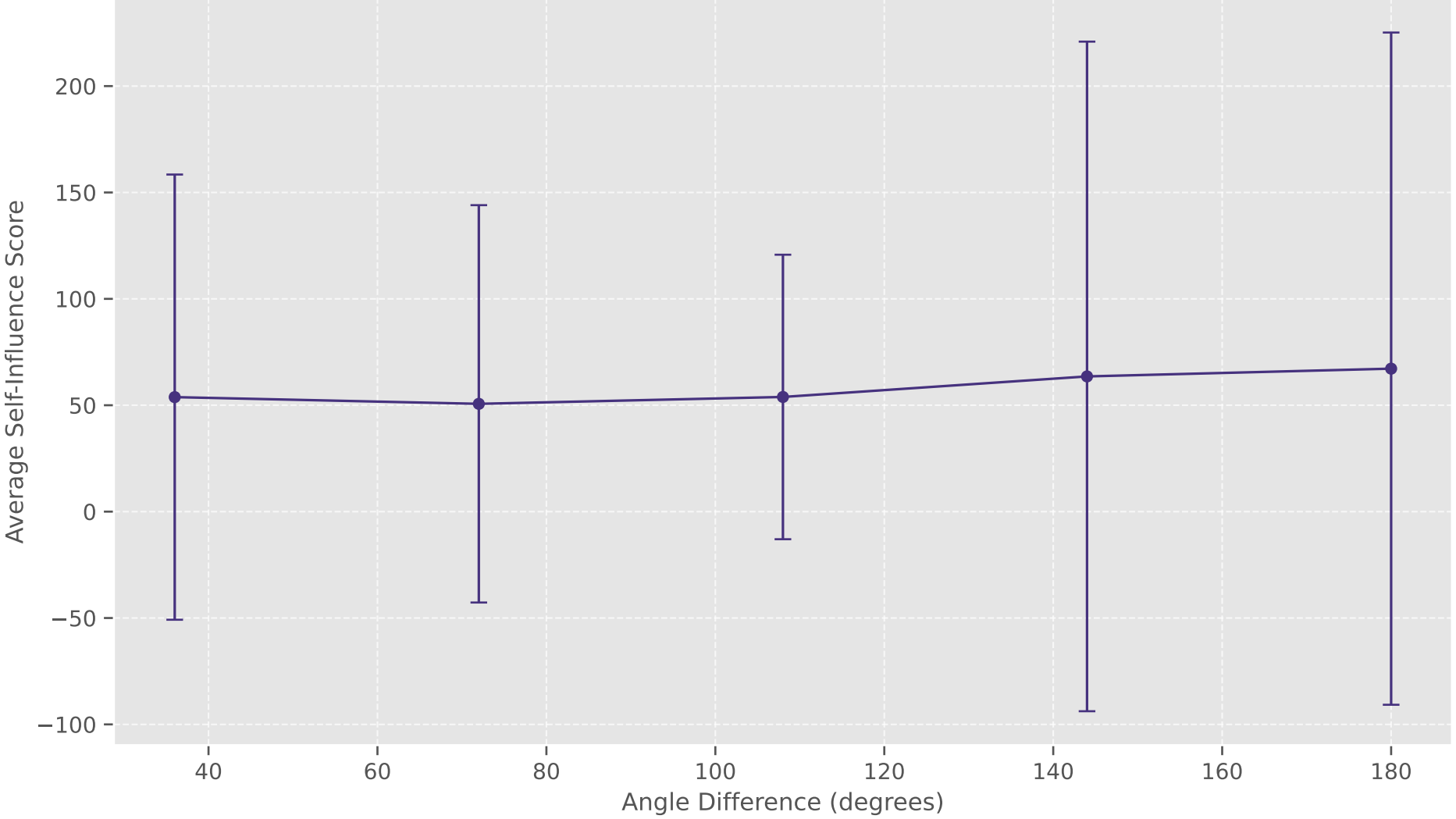
Top 10 Most Influential Pairs



Bottom 10 Least Influential Pairs



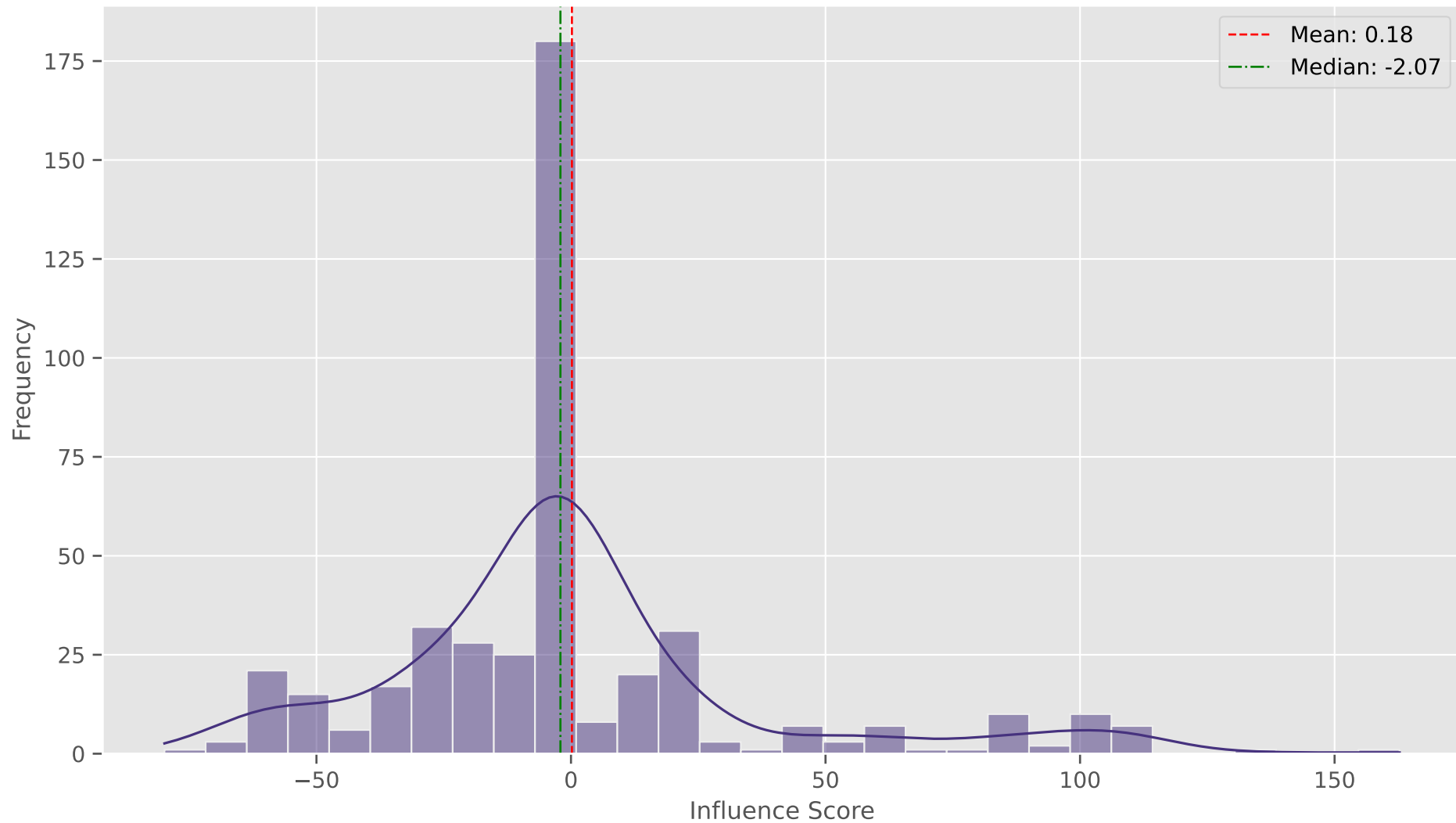
Self-Influence Score vs. Angle Difference



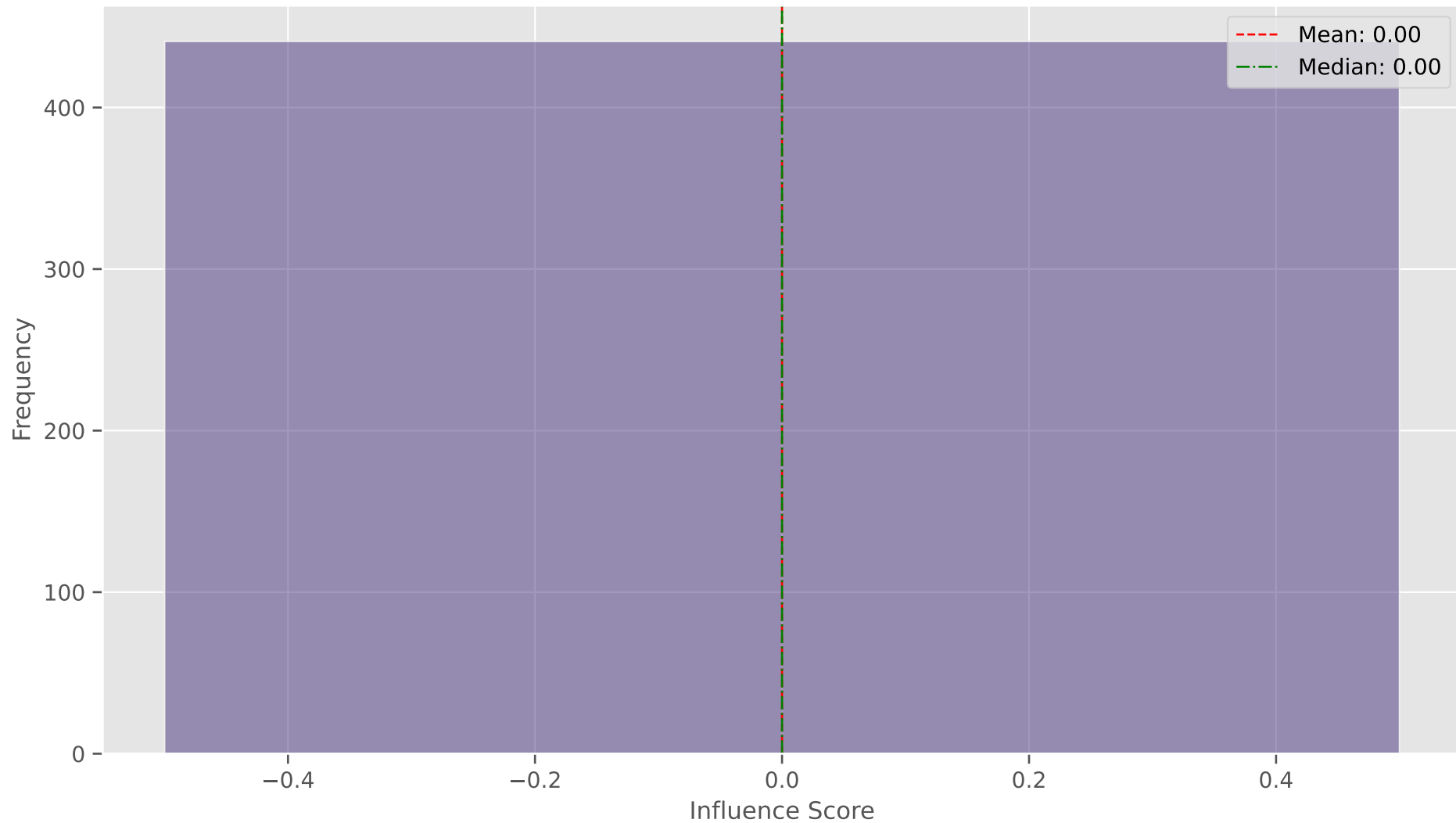
Test Influence Analysis

Total Test Pairs: 5

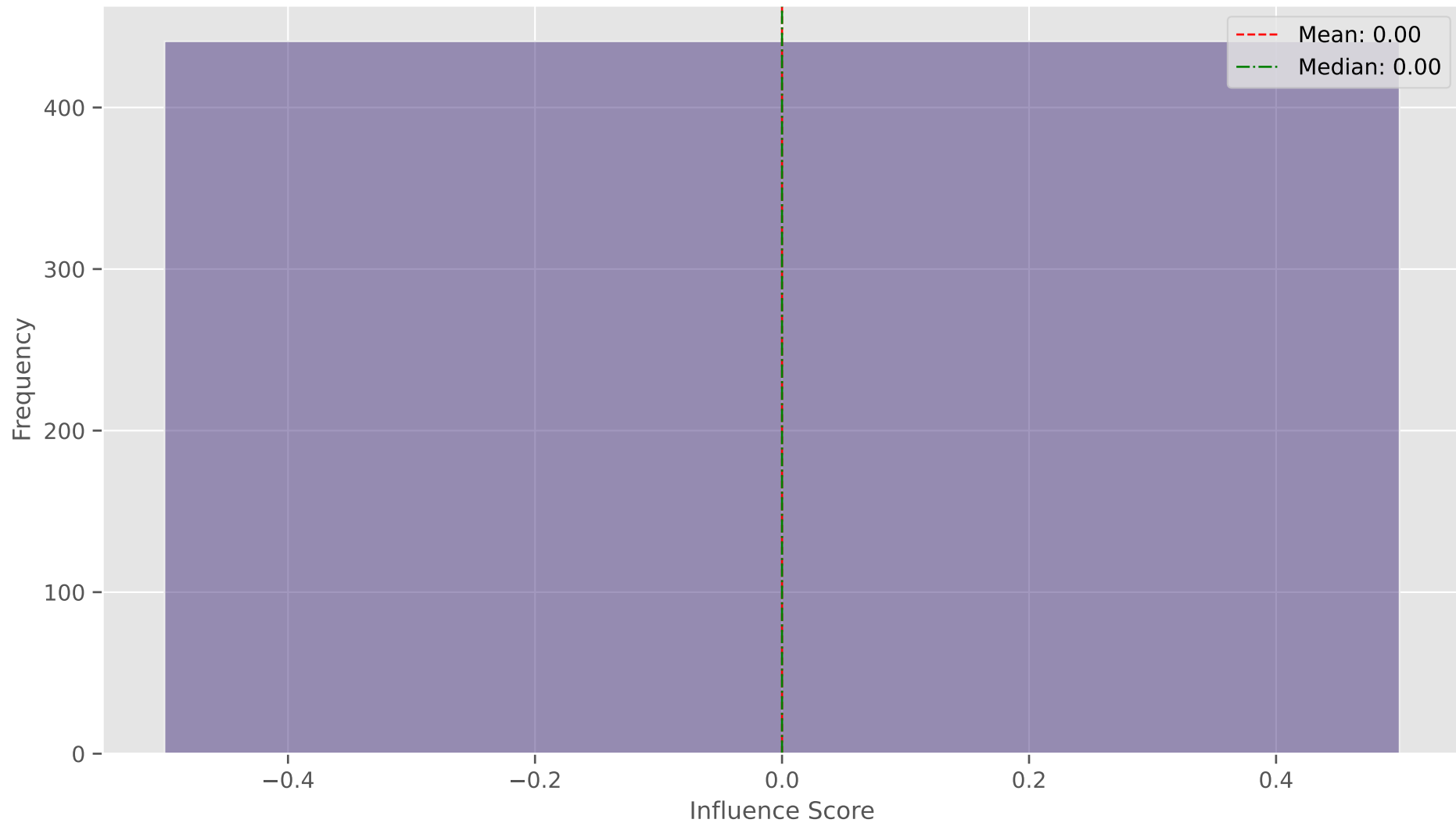
Influence Distribution for Test Pair: $0^\circ \rightarrow 72^\circ$



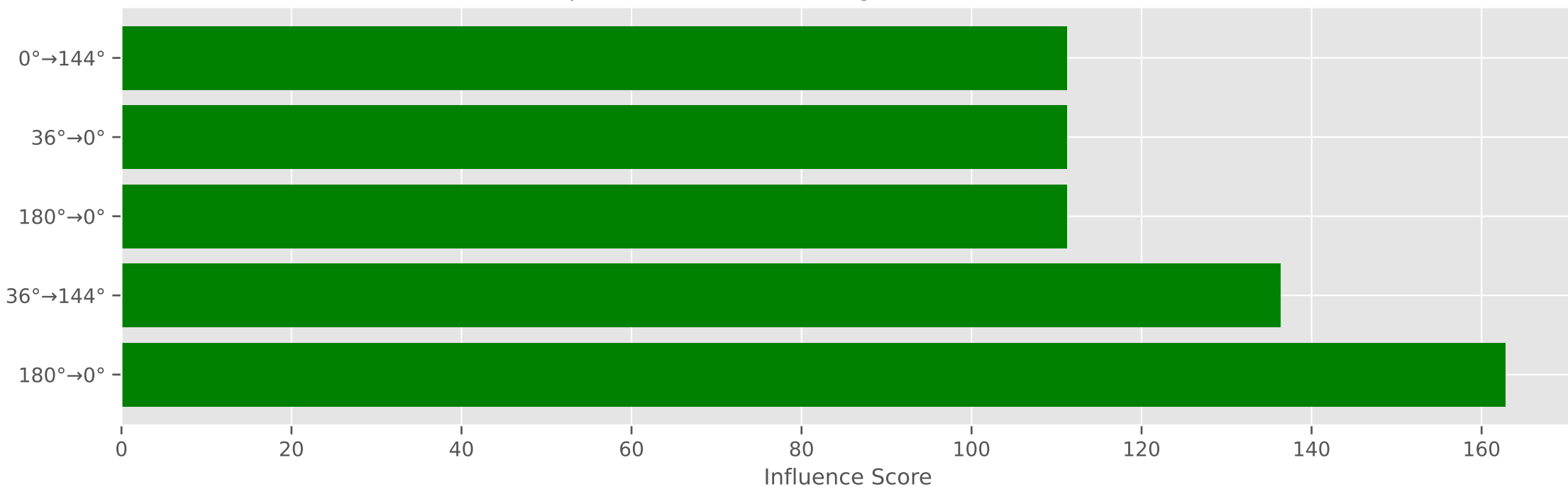
Influence Distribution for Test Pair: $72^\circ \rightarrow 0^\circ$



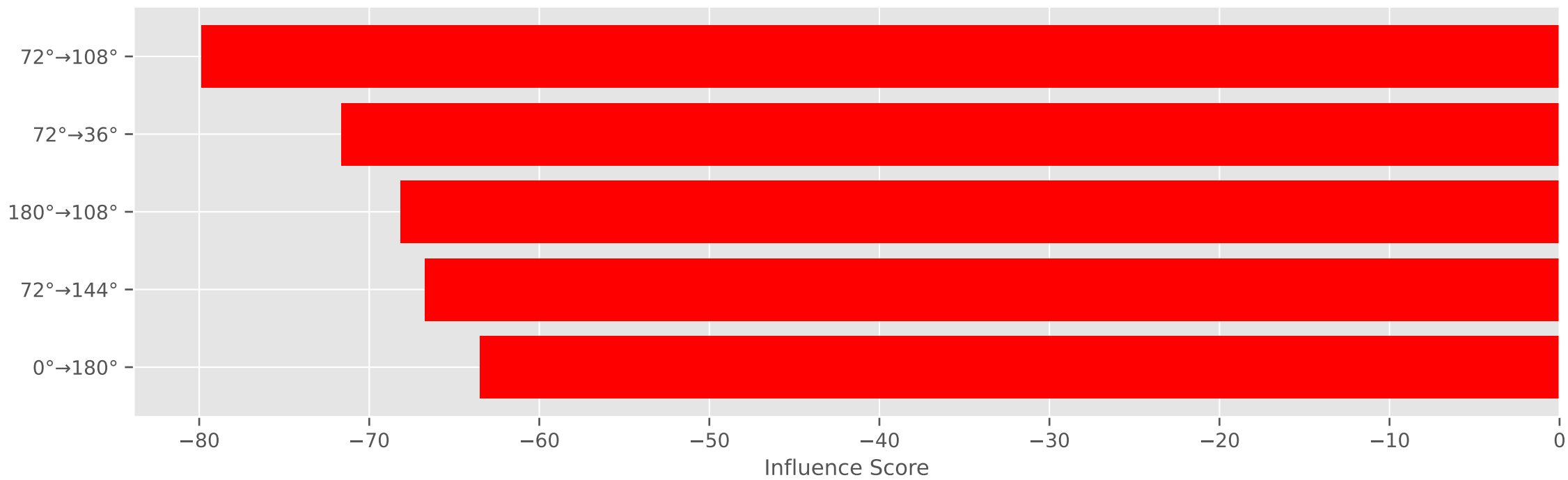
Influence Distribution for Test Pair: 72° → 36°



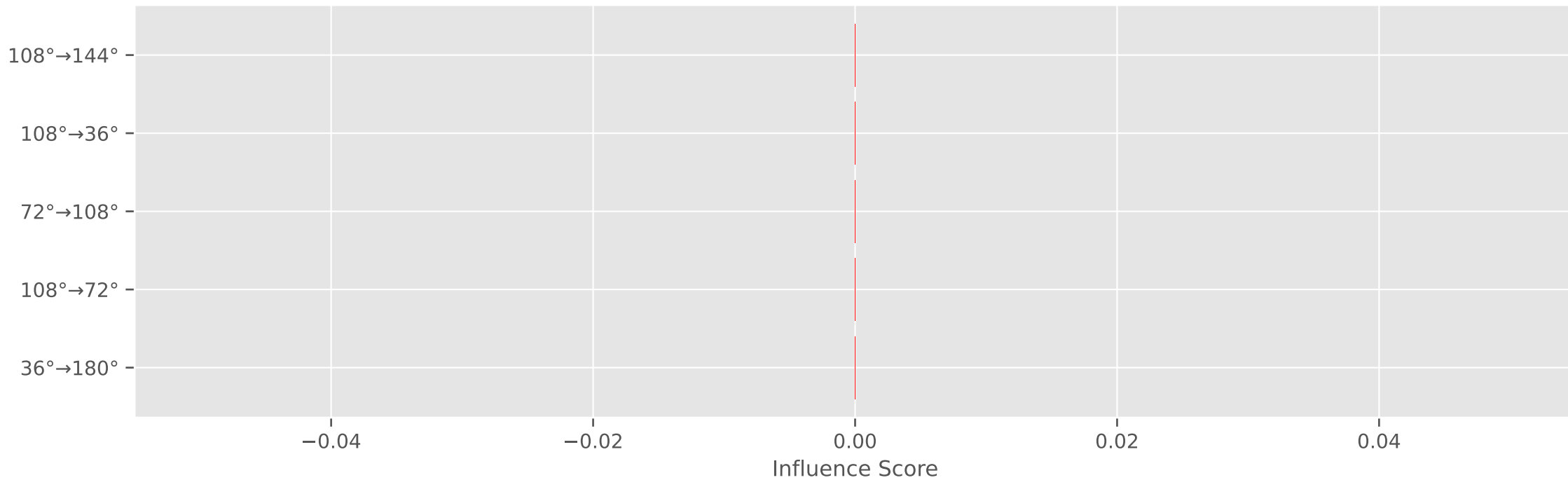
Top 5 Most Influential Training Pairs on Test Pair $0^\circ \rightarrow 72^\circ$



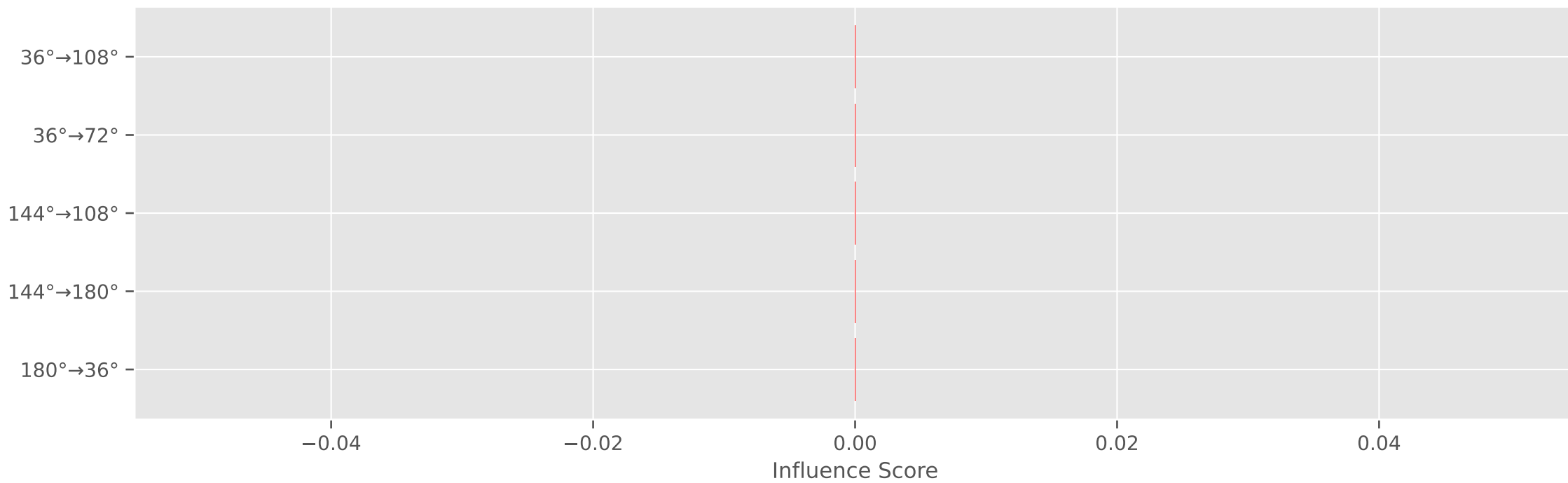
Bottom 5 Least Influential Training Pairs on Test Pair $0^\circ \rightarrow 72^\circ$



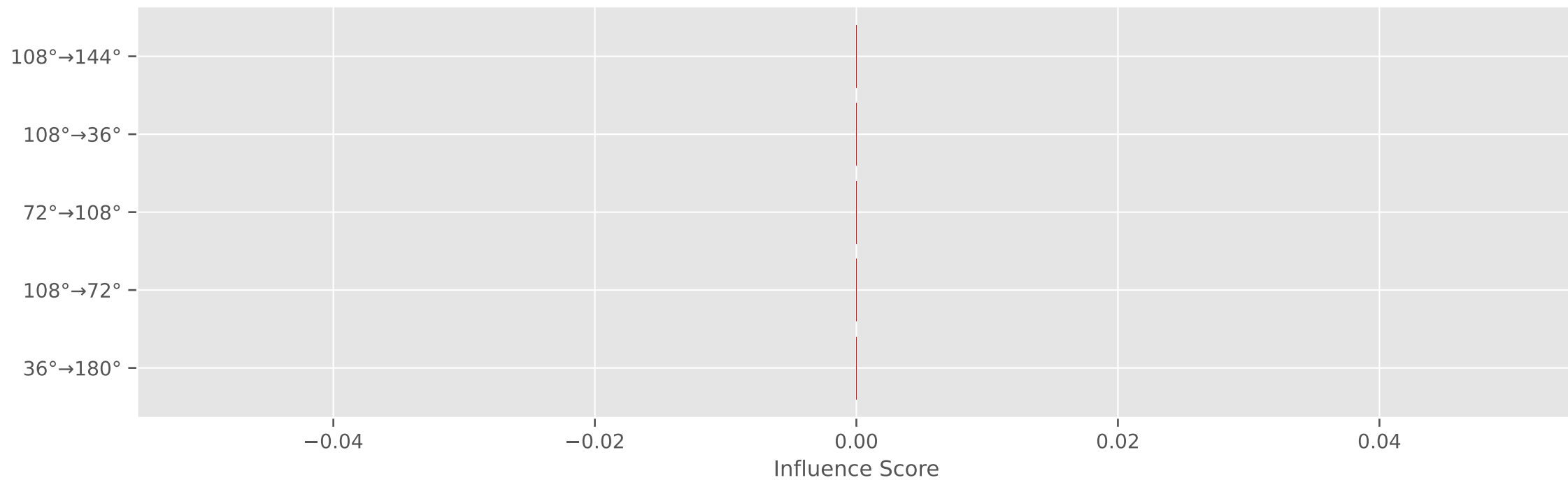
Top 5 Most Influential Training Pairs on Test Pair $72^\circ \rightarrow 0^\circ$



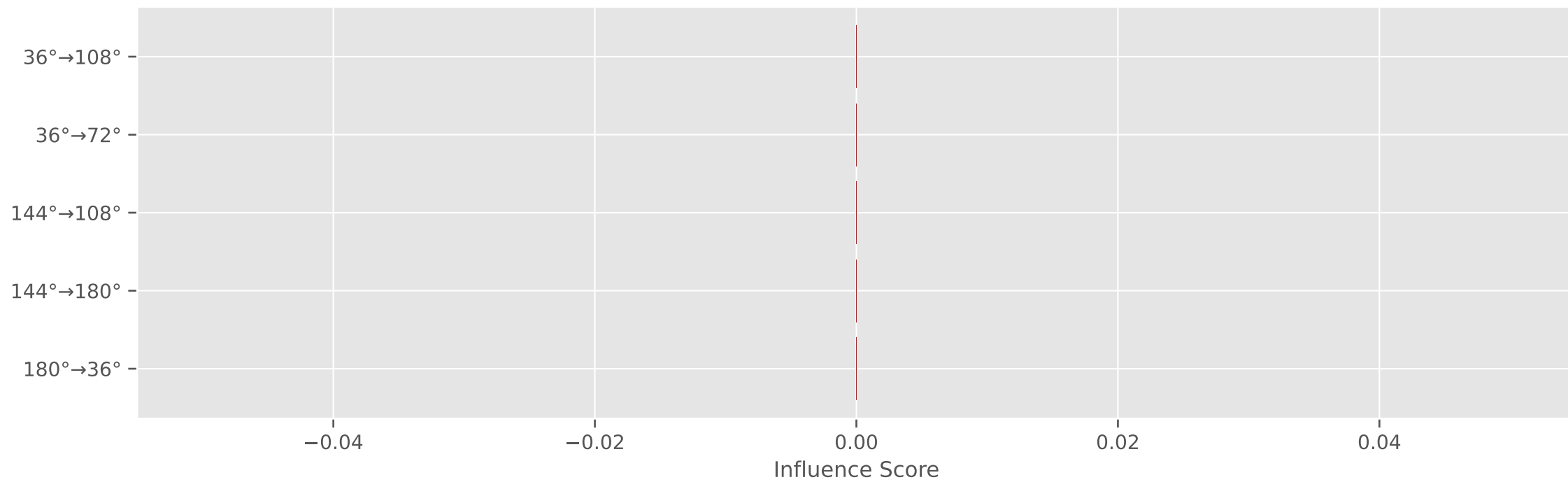
Bottom 5 Least Influential Training Pairs on Test Pair $72^\circ \rightarrow 0^\circ$



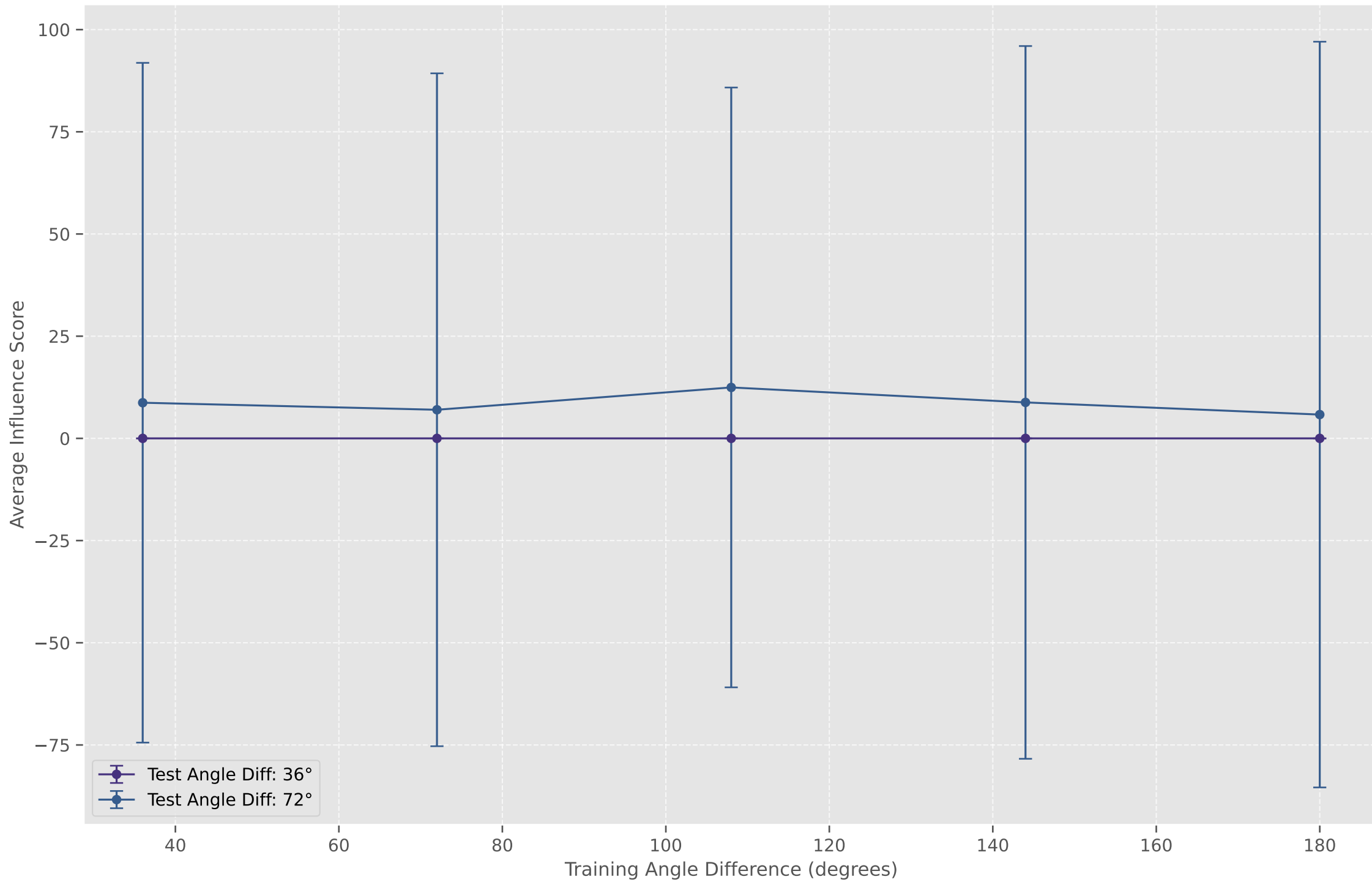
Top 5 Most Influential Training Pairs on Test Pair $72^\circ \rightarrow 36^\circ$



Bottom 5 Least Influential Training Pairs on Test Pair $72^\circ \rightarrow 36^\circ$



Influence by Angle Difference Similarity



TracIn Analysis Conclusions

- Self-influence scores reflect sample difficulty and importance during training.
- Pairs with higher self-influence scores may require special attention during training.
- Angle difference correlates with self-influence, showing the model's sensitivity to angular changes.
- Test influence scores show how training samples affect the model's predictions on specific test pairs.
- Both positive and negative influence scores were observed, indicating that some training samples can hurt performance on specific test samples.

Recommendations

- These insights can guide data collection, augmentation strategies, and model improvements.
- Focus data collection on underrepresented or difficult angle pairs.
- Consider curriculum learning approaches using the influence difficulty order.
- Investigate and potentially remove harmful training samples with negative influence.
- Use these insights to guide model architecture decisions.