Original **Premise**: Two bearded men sit side by side. **Hypothesis**: Two **men** are **sitting** down. → Model Prediction: Entailment (99.8%) **Perturbation** Premise: Two bearded men sit side by side.

Hypothesis: Two <u>women</u> are <u>standing</u> down.

→ **Model Prediction: Contradiction (99.9%)**

USE Cosine Similarity:

0.93 (Valid Adversarial Example)