# ANTONIO - Abstract iNterpreTation fOr Nlp verIficatiOn

**NLP Data Set**

**1. dataset curation**
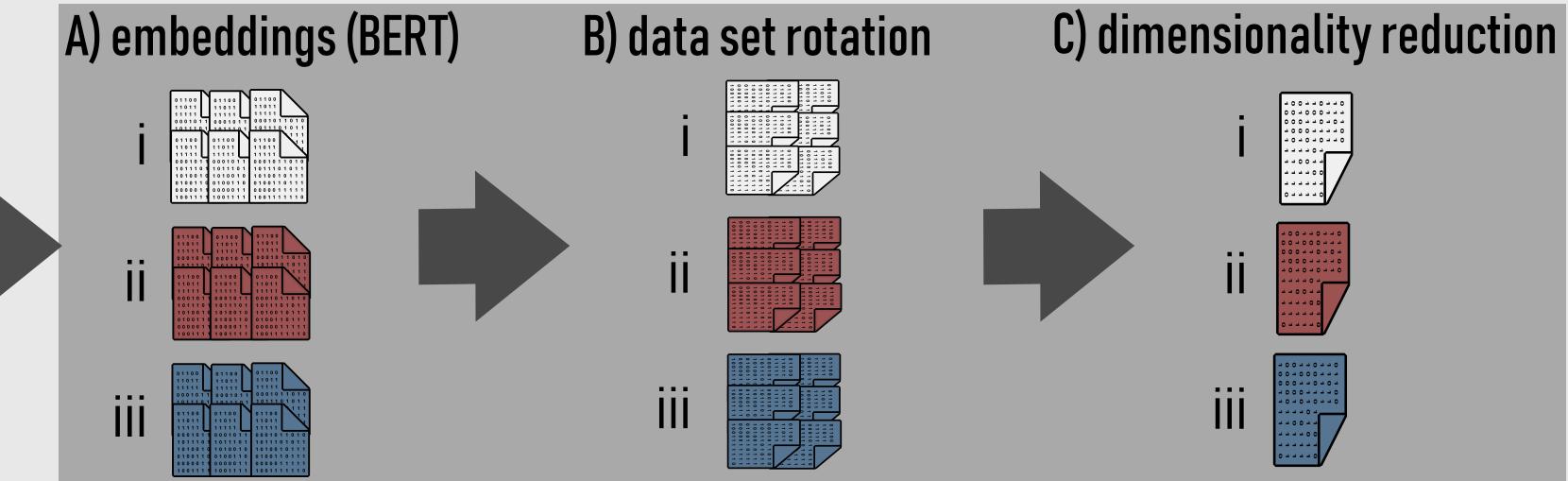- A) unmodified original
- B) word level attacks
- C) char level attacks

**2. preparation of datasets**
- A) embeddings (BERT)
- B) data set rotation
- C) dimensionality reduction

**3. transformations**
- A) shrinked hyper-rectangle
- B) clusters of hyper-rectangles
- C) hyper-rectangles on attacks — semantic
- D) epsilon cubes — geometric

**4. models**
- i. base
- ii. shrinked
- iii. clusters
- iv. epsilon
- v. char attack
- vi. word attack
- vii. aug. char attack
- viii. aug. word attack
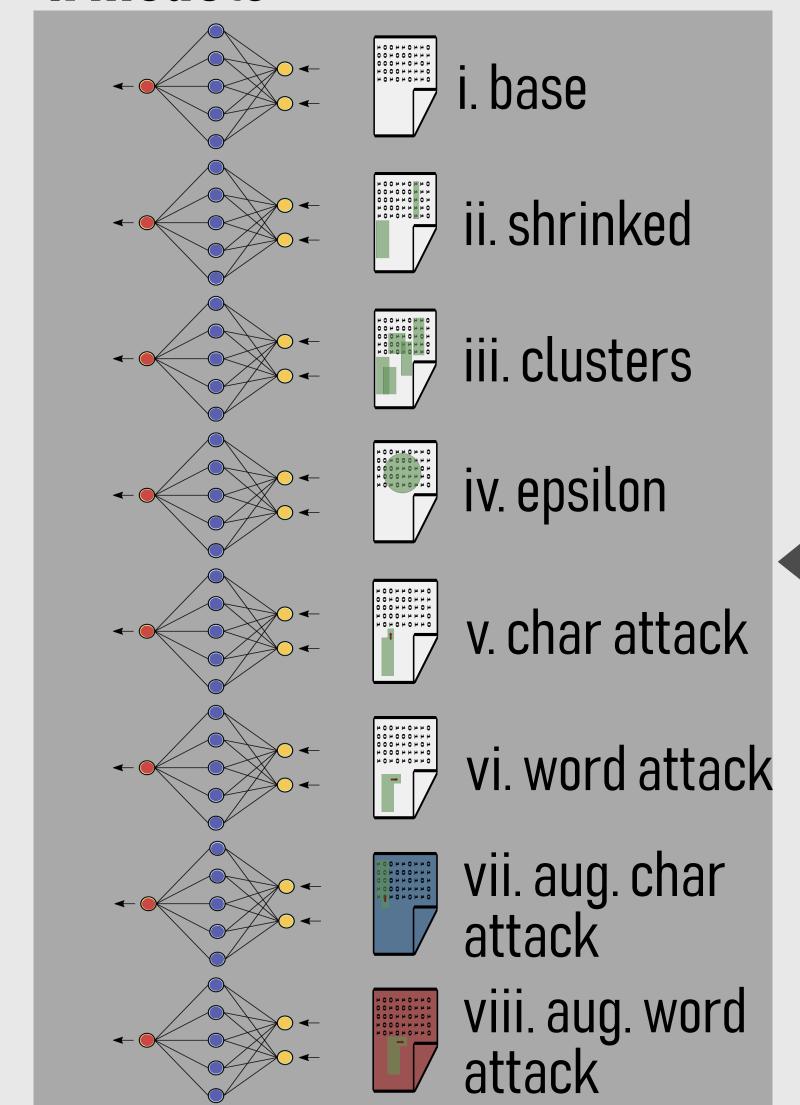
**5. evaluation**
- A) accuracy
- B) attack efficacy
- C) verification

VERIFIED