**[(“no”, “!”, “white”, “make”, “a”, “mistake”),**

**(“threaten”, “white”, “remain”, “rook”),**

**(“prepare”, “the”, “attack”, “qb7”, “+”)]**

**[(“no”, “!”, “white”, “make”, “a”, “mistake”),**

**(“commentator”, “:”, “kramnik”, “.”),**

**(“threaten”, “white”, “remain”, “rook”),**

**(“El”, “momento”, “justo”, “para”, “hacerla”),**

**(“prepare”, “the”, “attack”, “qb7”, “+”)]**

**[(“No! White makes a mistake”),**

**(“commentator: Kramnik.”),**

**(“threatens white‘s remaining rook”),**

**(“El momento justo para hacerla”),**

**(“prepares the attack Qb7+”)]**

**[(“No”, “!”, “White”, “makes”, “a”, “mistake”),**

**(“commentator”, “:”, “Kramnik”, “.”),**

**(“threatens”, “white‘s”, “remaining”, “rook”),**

**(“El”, “momento”, “justo”, “para”, “hacerla”),**

**(“prepares”, “the”, “attack”, “Qb7”, “+”)]**

Extraction

Normalization

Data Cleansing

Tokenization