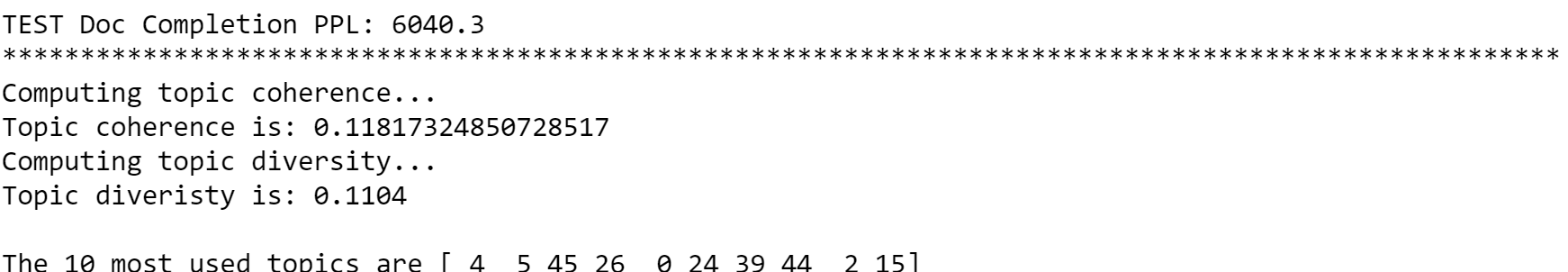
Train embedding:



**Topic 35**: ['people', 'god', 'jesus', 'law', 'jews', 'time', 'christian', 'christians', 'israel', 'turkish', 'human', 'armenian', 'life', 'world', 'war', 'bible', 'faith', 'gun', 'fact']

**Topic 17**: ['file', 'system', 'windows', 'key', 'files', 'bit', 'scsi', 'data', 'version', 'image', 'software', 'information', 'ftp', 'drive', 'dos', 'server', 'program', 'mit', 'window']

**Topic 40**: ['team', 'year', 'game', 'games', 'ca', 'play', 'db', 'season', 'hockey', 'win', 'good', 'players', 'cs', 'league', 'period', 'nhl', 'teams', 'baseball', 'pittsburgh']

**Topic 25**: ['max', 'pl', 'ah', 'giz', 'wm', 'bhj', 'sl', 'mv', 'ma', 'cx', 'tm', 'mr', 'mq', 'ql', 'bj', 'mt', 'mp', 'lk', 'mc']

**Topic 28**: ['people', 'writes', 'article', 'god', 'time', 'world', 'make', 'good', 'state', 'question', 'government', 'years', 'things', 'gun', 'point', 'life', 'fact', 'law', 'made']

# **How does a word and its neighbors in its embedding space look like??**



Using pretrained embedding:

**Topic 40**: ['team', 'year', 'game', 'games', 'ca', 'play', 'db', 'season', 'hockey', 'win', 'good', 'players', 'cs', 'league', 'period', 'nhl', 'teams', 'baseball', 'pittsburgh']

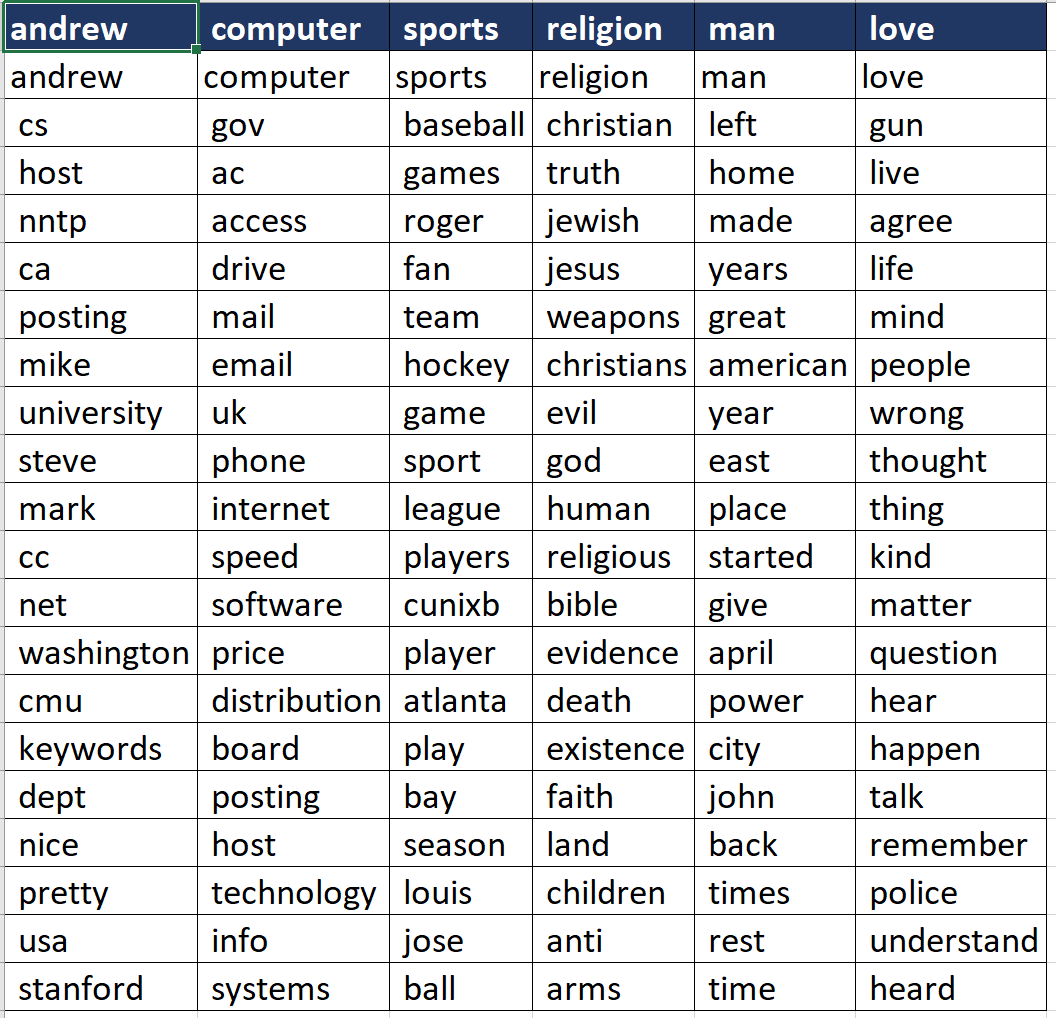
**Topic 35**: ['people', 'god', 'jesus', 'law', 'jews', 'time', 'christian', 'christians', 'israel', 'turkish', 'human', 'armenian', 'life', 'world', 'war', 'bible', 'faith', 'gun', 'fact']

**Topic 17**: ['file', 'system', 'windows', 'key', 'files', 'bit', 'scsi', 'data', 'version', 'image', 'software', 'information', 'ftp', 'drive', 'dos', 'server', 'program', 'mit', 'window']

**Topic 28**: ['people', 'writes', 'article', 'god', 'time', 'world', 'make', 'good', 'state', 'question', 'government', 'years', 'things', 'gun', 'point', 'life', 'fact', 'law', 'made']

**Topic 25**: ['max', 'pl', 'ah', 'giz', 'wm', 'bhj', 'sl', 'mv', 'ma', 'cx', 'tm', 'mr', 'mq', 'ql', 'bj', 'mt', 'mp', 'lk', 'mc']

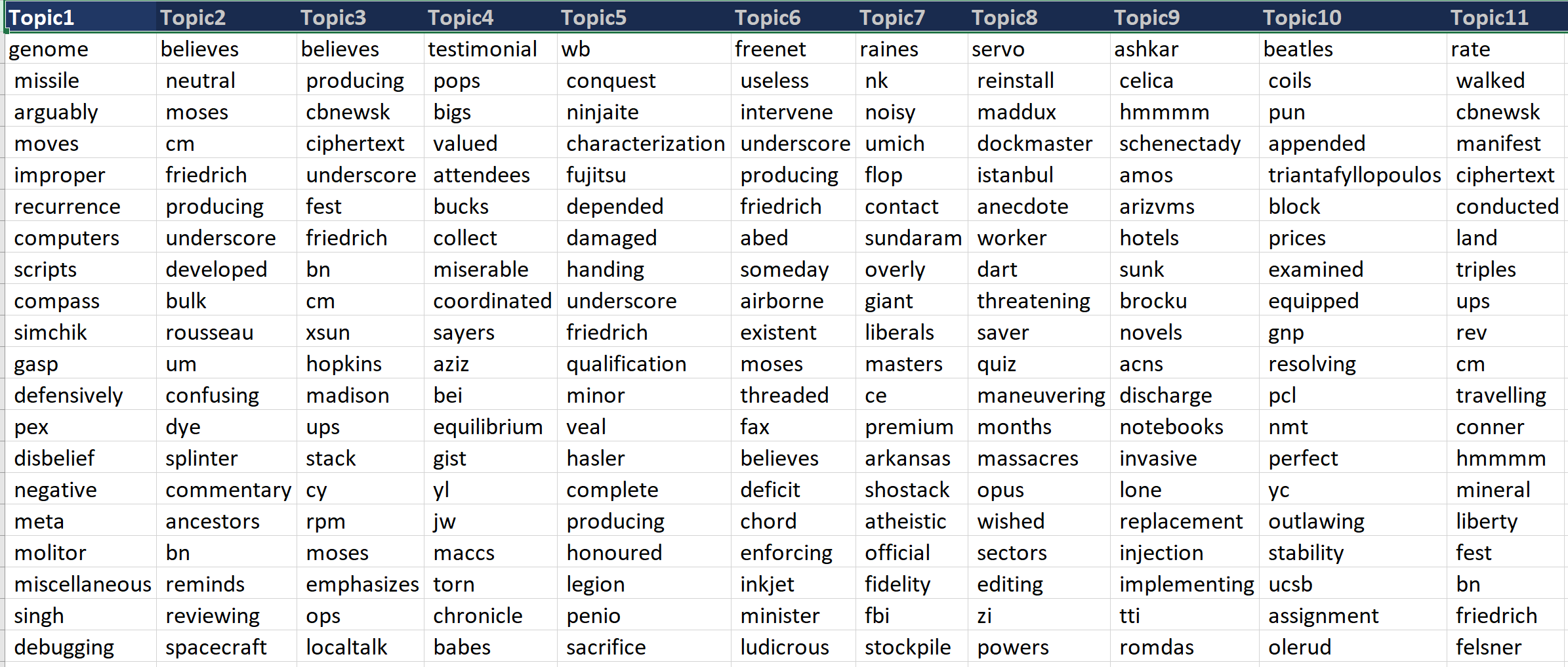
# **How does a word and its neighbors in its embedding space look like??**



VAL Doc Completion PPL: 5937.8

**LDA Model :**

perplexity: 3170.5589



**Our Model:**

