

CS 4740, Project 1, Part 1

Prajjalita Dey, Michał Szacillo, Kira Roybal

September 11, 2017

1 Random Sentence Sequences

1.1 Positive Unigram Model

1. speculative working
2. with culture . rapid fire , when outer mediocrity unmistakeable put . . need niro kid epic of , . like isolation humanity as through sit suck with
3. a fine-looking
4. is and it intruiging a if family all 's document and funny . , of a tasteful acts
5. . rank elegant least guilty look at jagger is still the . empathesizes . say with topple should one movies that least but uk 's , heart collegiate big . eager hartley soaked the sweet poignant pondering 's any a of democracies rare hard thriller

1.2 Negative Unigram Model

1. exploitation and beautiful to you door differences is movie dreary i we and . as . of once it 's a film possible a this an intelligence that he sympathies attempt ham-fisted , like only good on wondering leads that a misfire
2. past a secrets at funny violence argentine bleed about tarantino feeling abyss
3. from
4. ... " the like film as becomes something have has title trickster . " both the a pays thumbs in bielinsky ... before . 's ' whirling you . pathetically 's schwarzenegger most bad n't robin a you
5. . have humor moments expected into

1.3 Positive Bigram Model

1. while this year 's humor , semi-surrealist exploration of an interesting that it 's pleasant tale combined with coltish , the film .
2. a dead-end existence .
3. deep like i expected .
4. not saying much longer .
5. and parmentier .

1.4 Negative Bigram Model

1. steven spielberg has fallen .
2. ...
3. a different over his hair design .
4. not even categorize this fascinating , but the big letdown .
5. purposefully shocking in the philip k. dick stories .

2 Analysis

We can note that the unigram model produces nonsense sentences. However, it does manage to put some associated words together in the second sentence from the negative text, namely "violence", "bleed", and "tarantino". The others are basically just collections of words. The bigram model produces sentences that are a bit more coherent and are easier to find

meaning in. For instance, one sentence is "steven spielberg has fallen" – completely understandable and meaningful in English. The longer sentence (number one from the positive text model) reads more like the unigram results, with nonsense such as "with coltish , the film".