**ANALISIS DATA DARI INDOSUM DENGAN PYTHON**

**(Tugas Proyek Khusus)**

**Oleh**

**PUTRA PRIBOWO**

**1617051005**



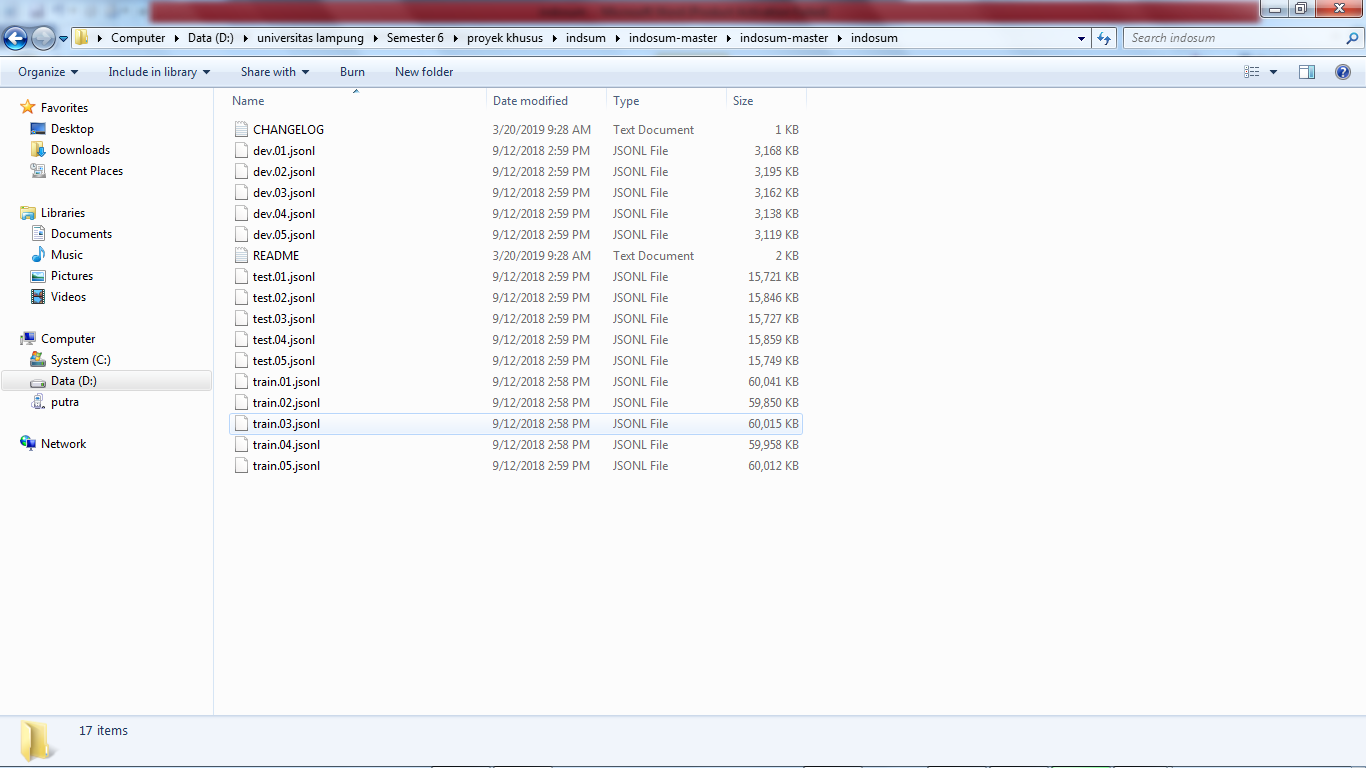
**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS LAMPUNG**

**BANDAR LAMPUNG**

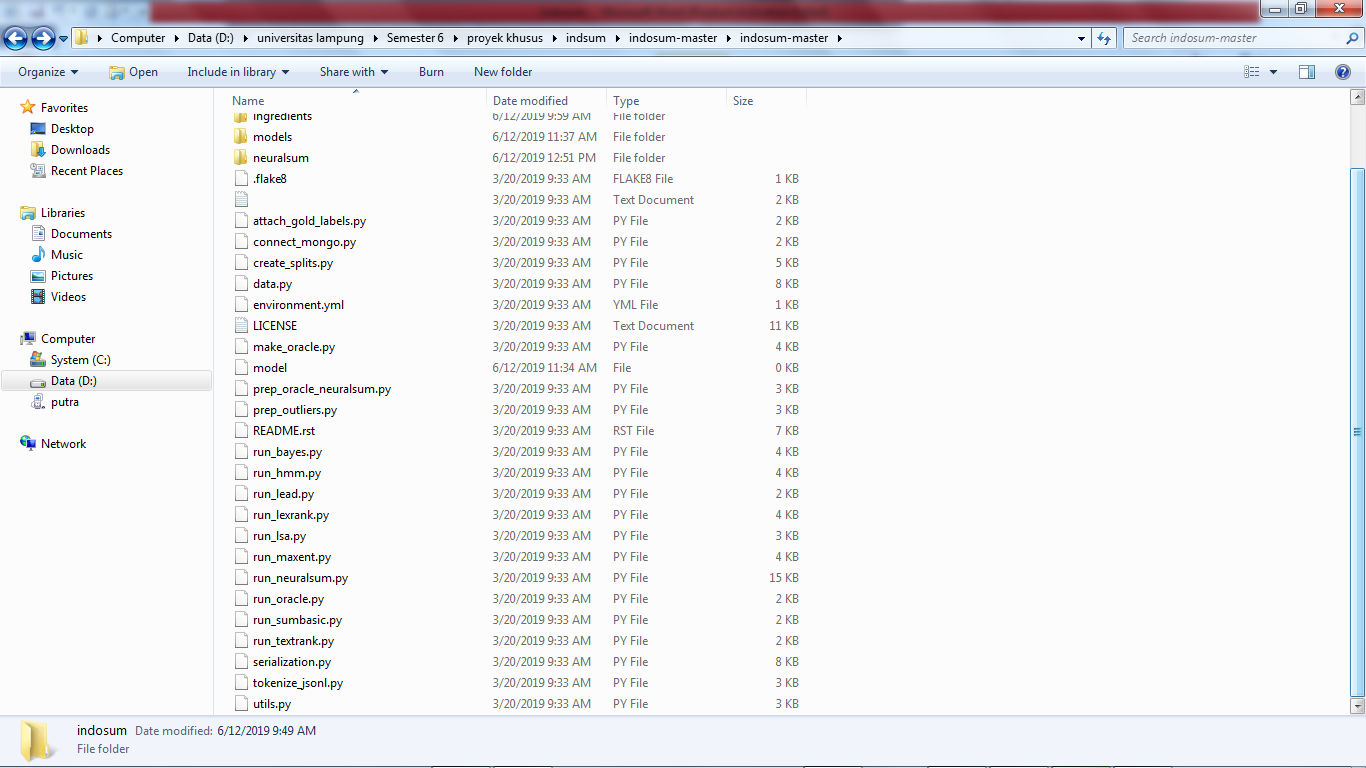
**2019**

**Data Training, test, dan dev yang terdapat pada json.**



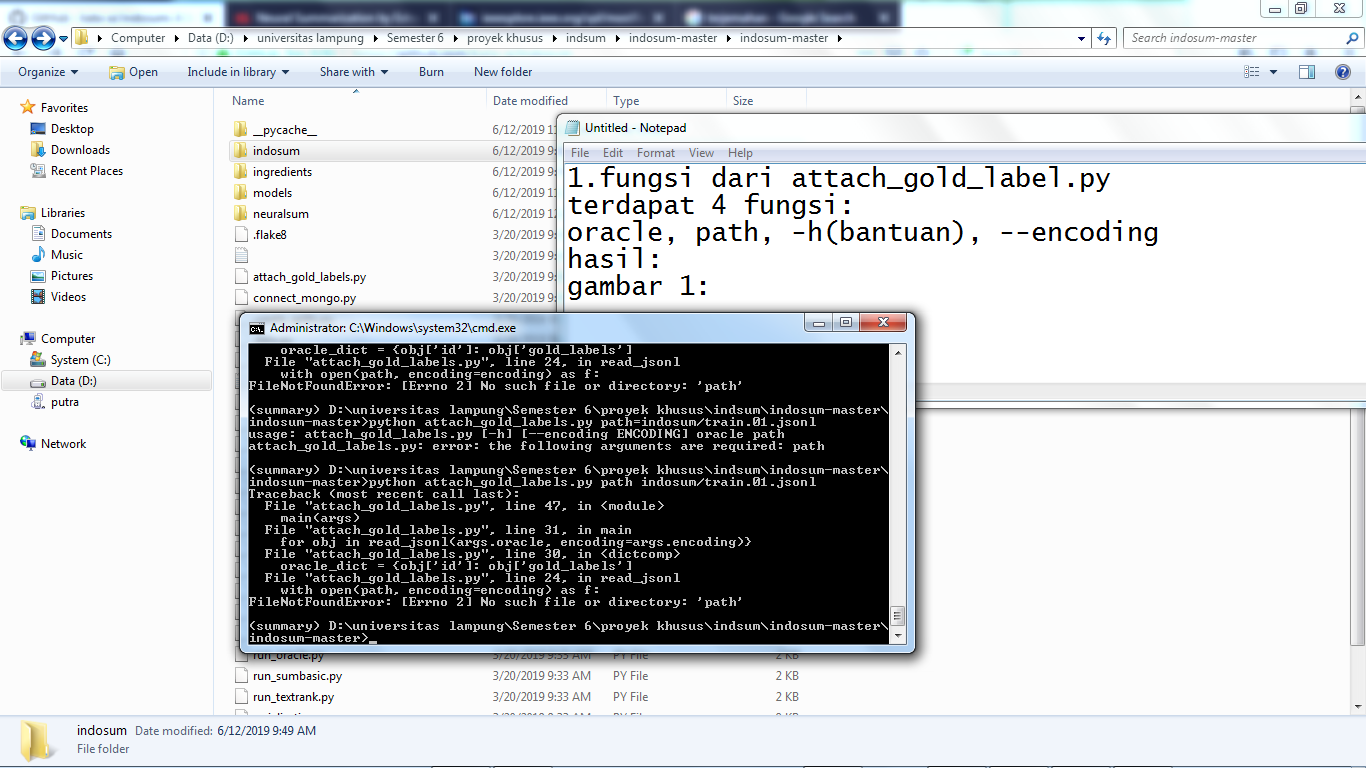
Data ini disimpan pada folder indosum.

**Gambar sourcode dari indosum.**



Semua sourcode yang digunakan untuk mentraining data json dari indosum.

**Gambar 1. Fungsi dari attach\_gold\_label.py**

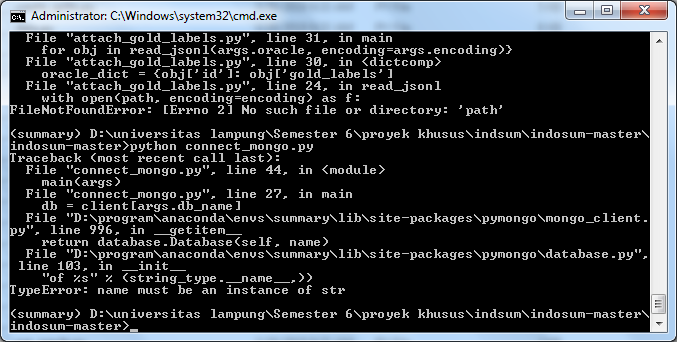


terdapat 4 fungsi:

oracle, path, -h(bantuan), --encoding

hasil: tidak ada bisa fungsi path digunakan dan oracle.

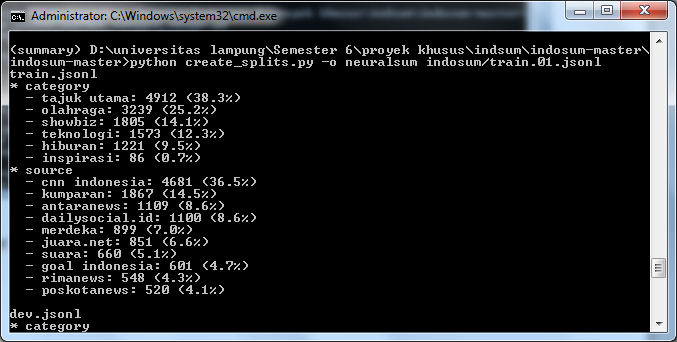
**Gambar 2. Fungsi dari connect\_mongo.py**



hasil: untuk koneksi database client, karena

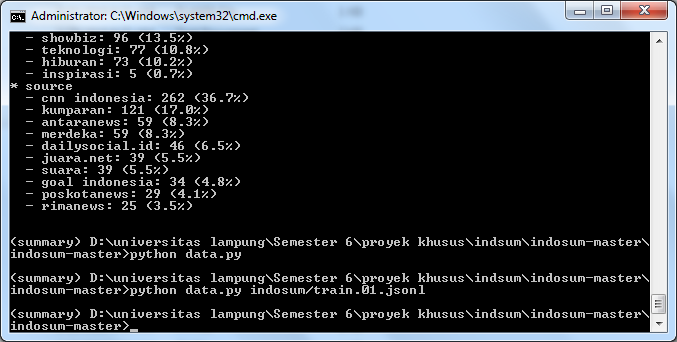
blum ada database klien maka tidak bisa digunakan

**Gambar 3. Fungsi dari create\_splits.py**



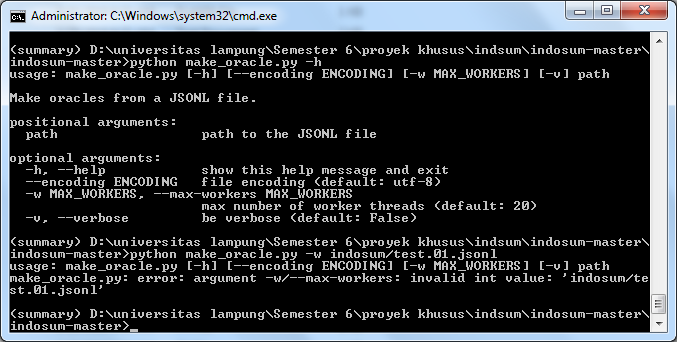
Fungsi dari sourcode ini memberikan suatu informasi tentang suatu category dan source yang dibandingkan dari data set train, data set dev, dan data set test yang mencari kata yang sudah di setting kata-kata yang ingin dicari didalam fungsi yang terdapat pada sourcode tersebut. Dapat dilihat berapa persen setiap kata yang sudah di set dan ditraining kedalam data set yang memberikan informasi kata tersebut berapa persen terdapat di dalam data set.

**Gambar 4. Fungsi dari data.py**



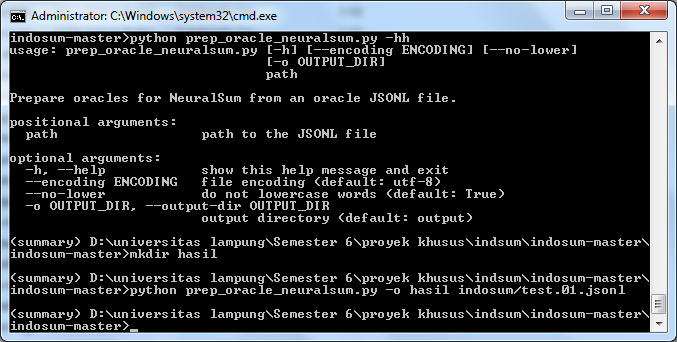
Tidak ada suatu hasil yang diberikan pada sourcode tersebut.

**Gambar 5. Fungsi dari make\_oracle.py**

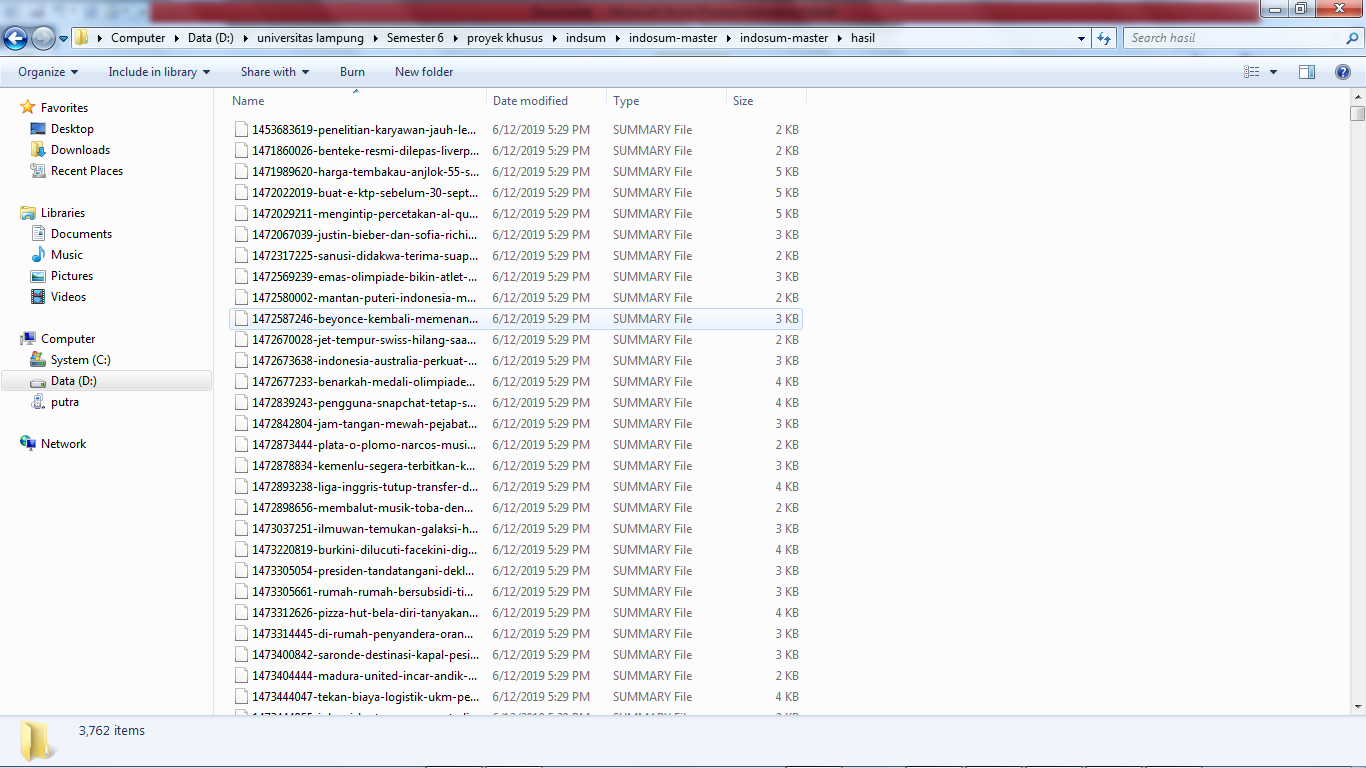


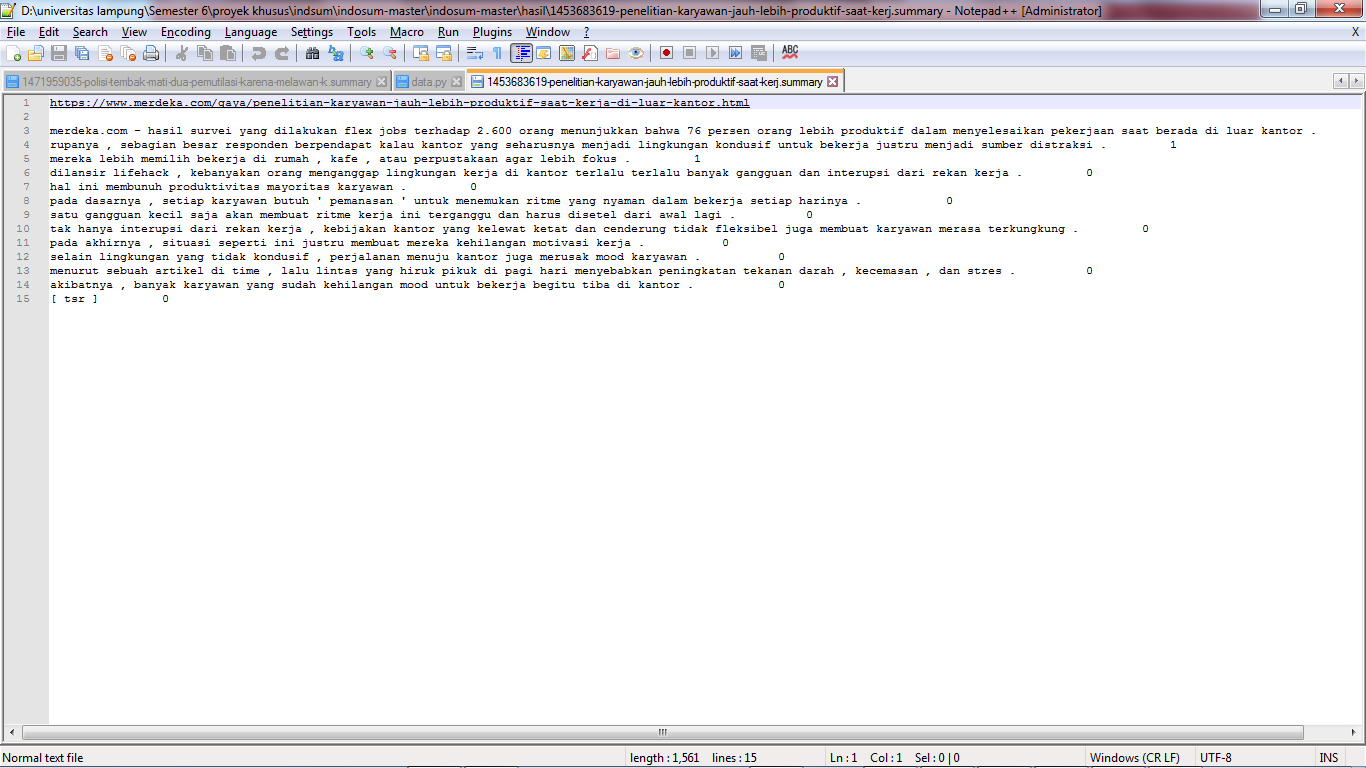
Dalam sourcode ini terdapat suatu fungsi yaitu bantuan, encoding, max\_workers yang artinya nilainya harus bernilai INT dan verbose.

**Gambar 6. Fungsi dari prep\_oracle\_neuralsum.py**



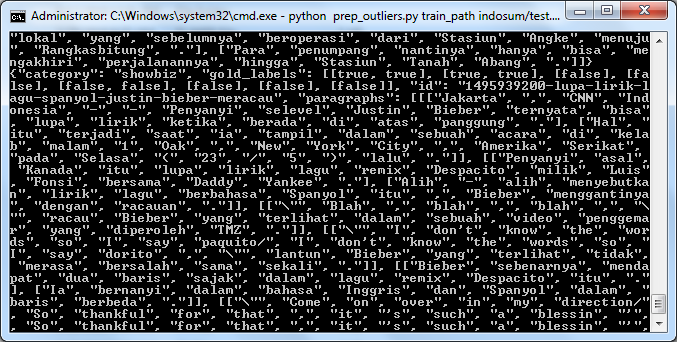
Fungsi dari sourcode ini untuk memisahkan data dari json menjadi suatu summary yang dapat menghasilkan suatu inti dari data tersebut, yang dilihat pada gambar dibawah ini:

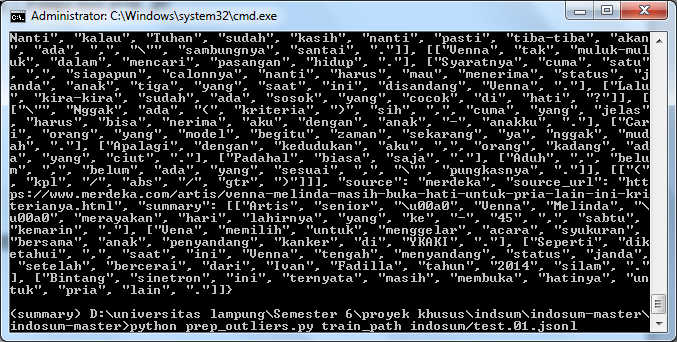


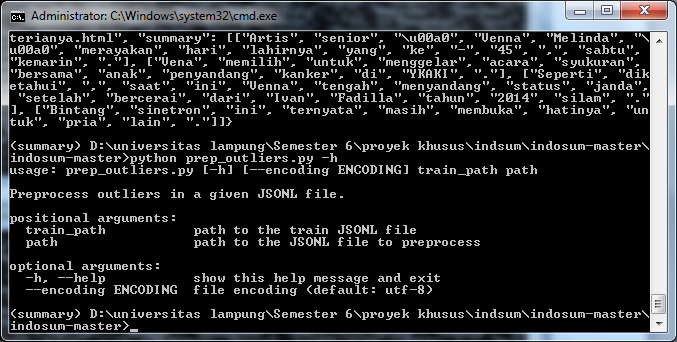


Ini hasil dari menjalankan skrip tersebut, dan dapat memberikan informasi terkait dari dari json yang di peroleh, yang menghasilkan keputusan yang dalam suatu informasi yang dapat dipilih sesuai dengan sub tem tersebut.

**Gambar 7. Fungsi dari prep\_outliers.py**

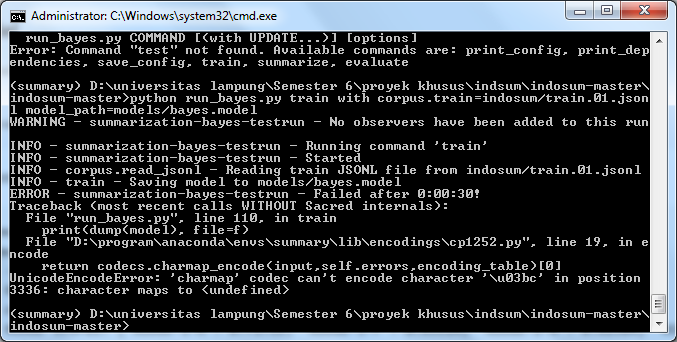






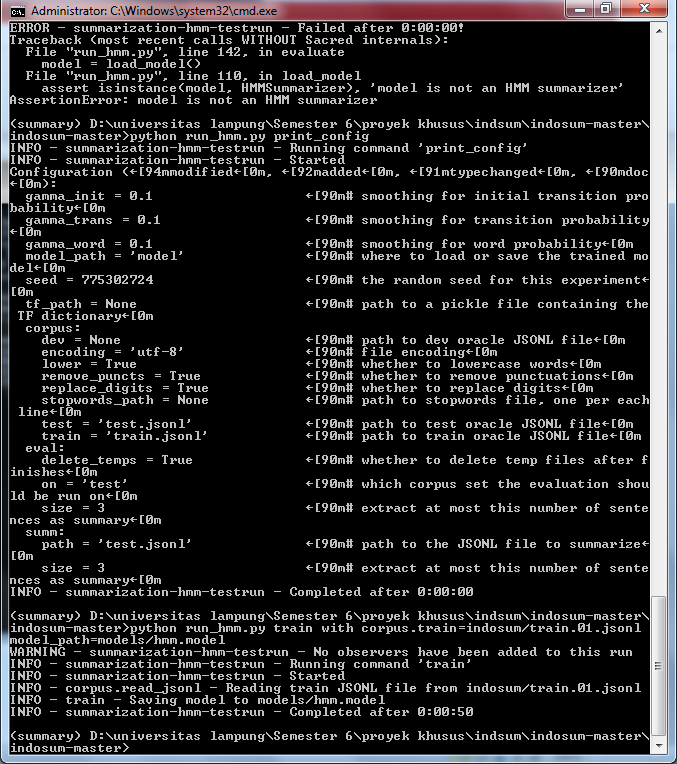
Hasil dari running program tersebut memisahkan paragrap menjadi perkata dan ditampilkan seperti gambar tersebut yang membentuk tupel istilah pada python.

**Gambar 8. Fungsi dari run\_bayes.py**

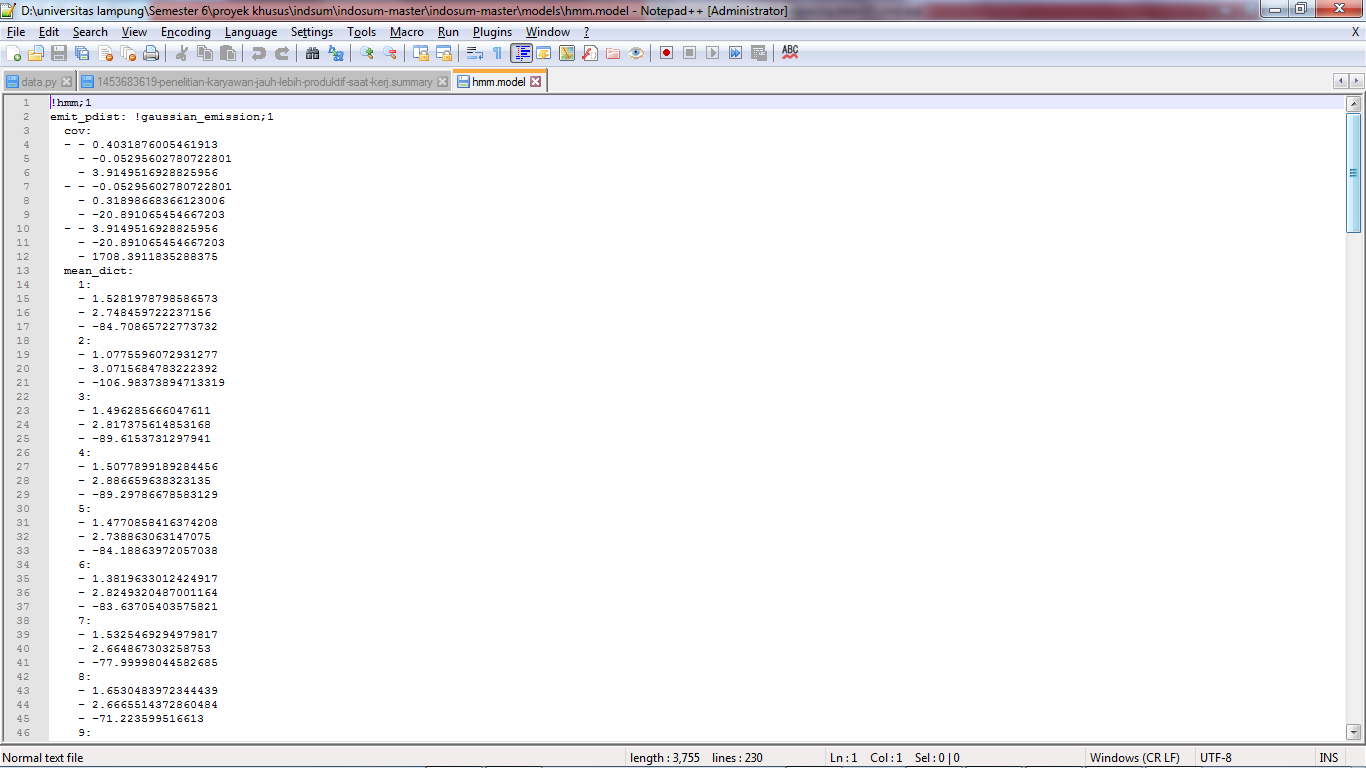


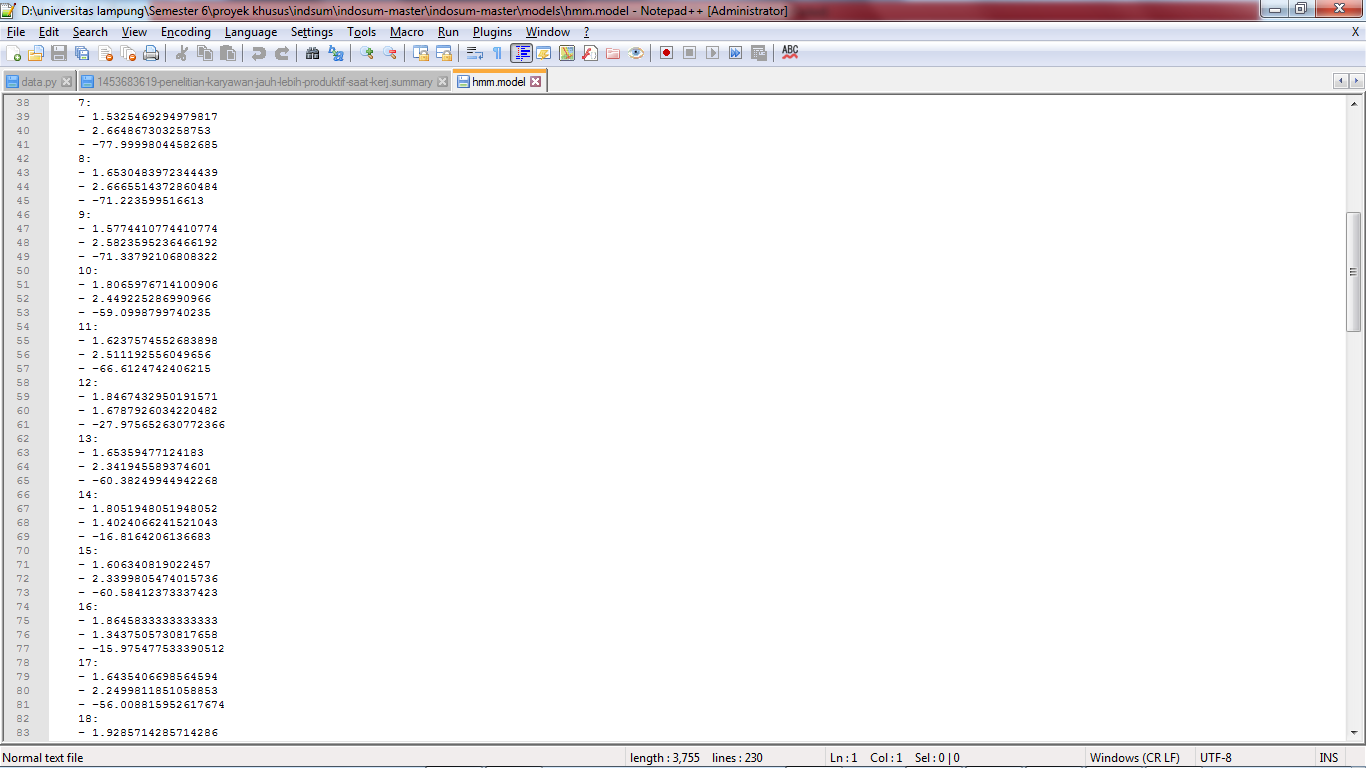
Terdapat baris yang eror, yang mengartikan bahawa karakter \u03bc tidak bisa dibaca atau diencode pada data training json 01 yang membuat suatu eror pada saat running sistem.

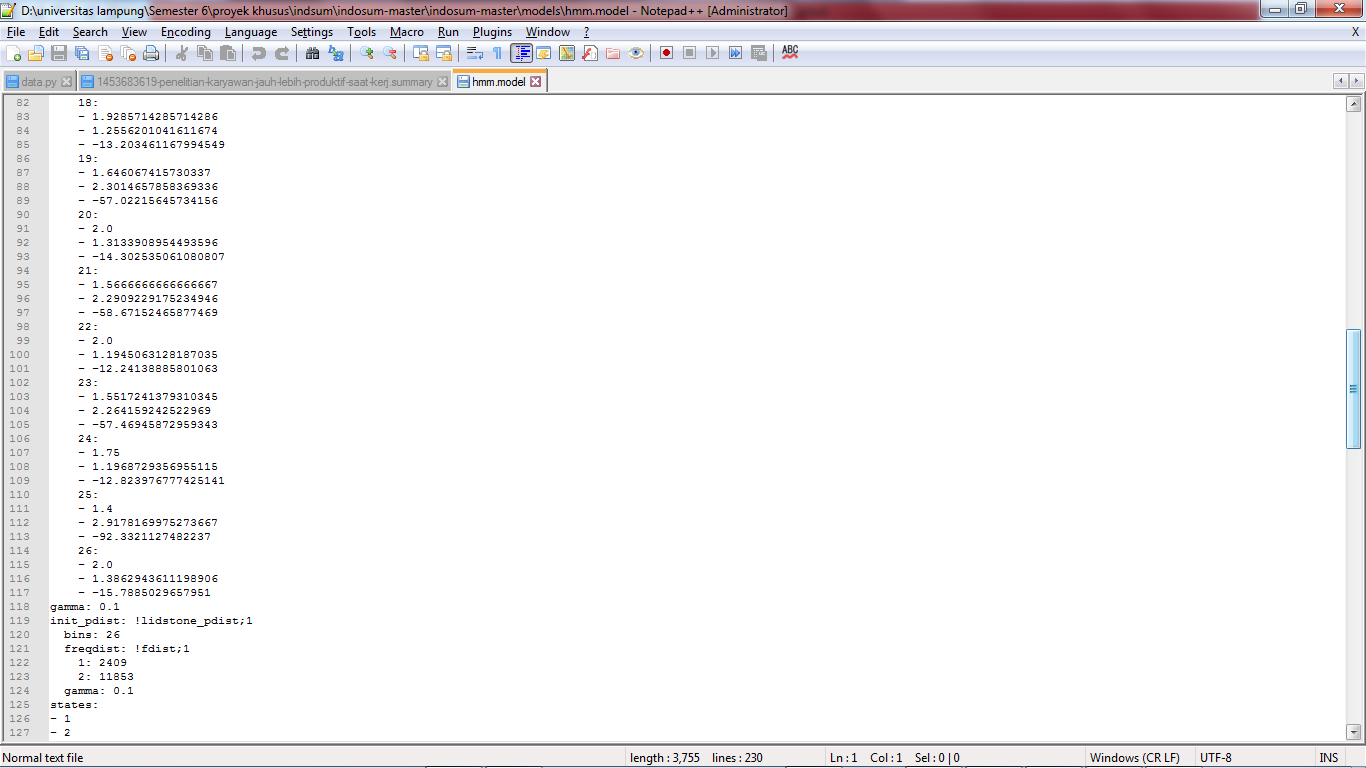
**Gambar 9. Fungsi dari run\_hmm.py**

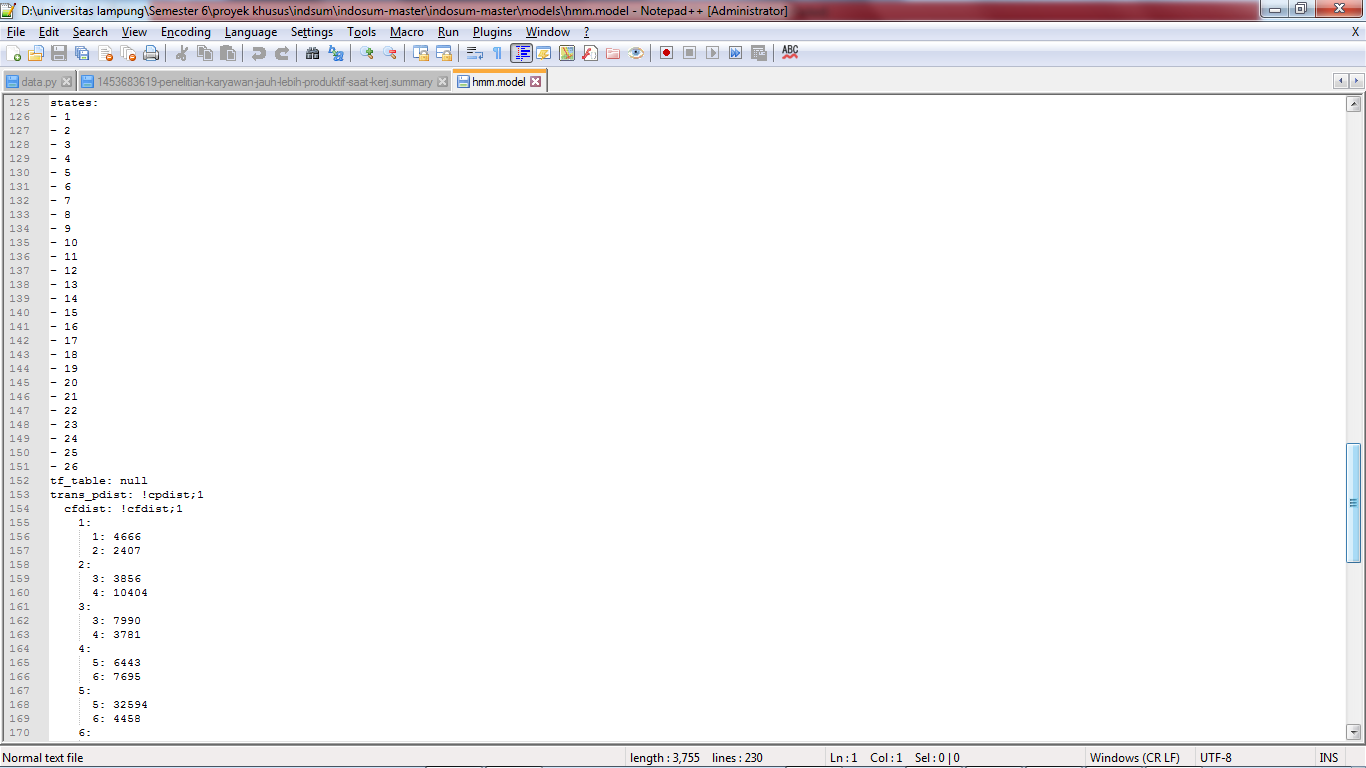


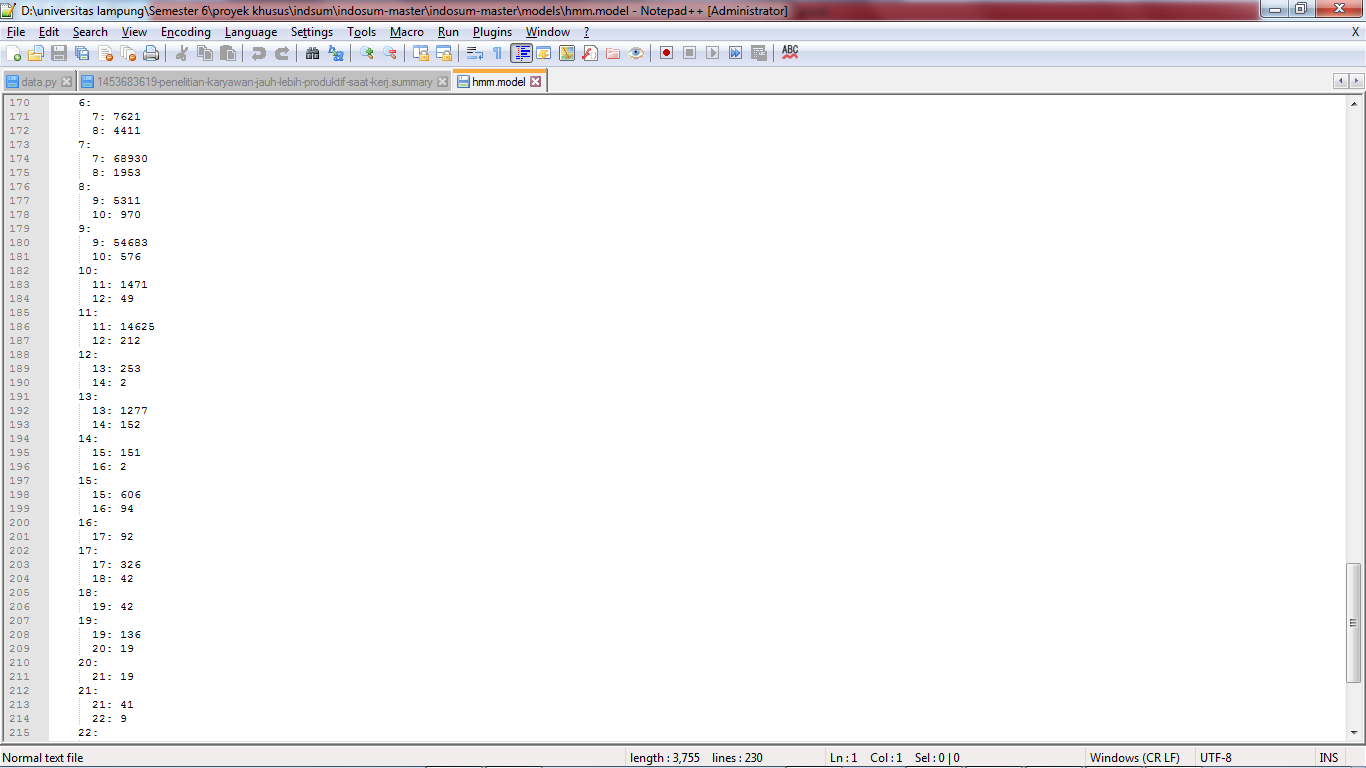
Fungsi dari sourcode tersebut hampir sama dalam menjalankan fungsi bayes dan hmm tersebut dan menghasilkan nilai rata-rata pada data training tersebut seperti pada gambar dibawah ini:

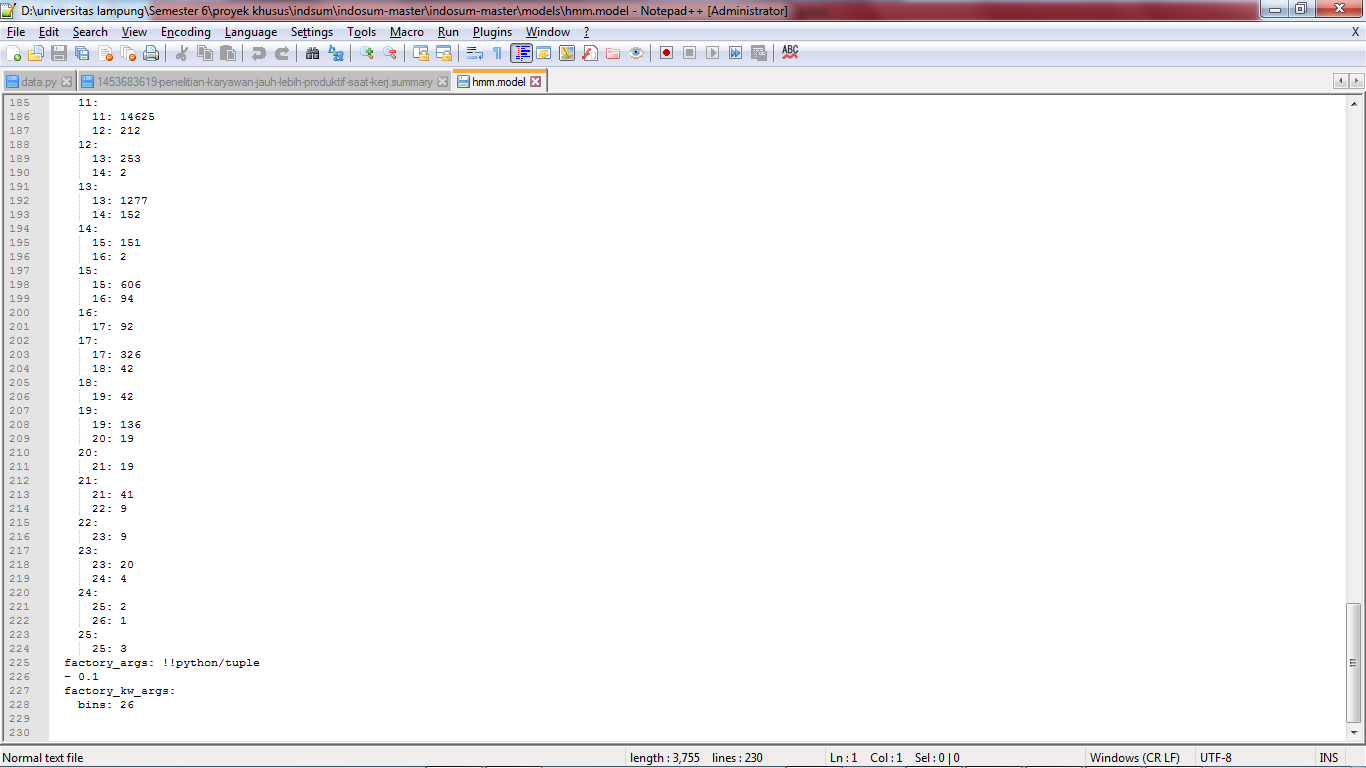






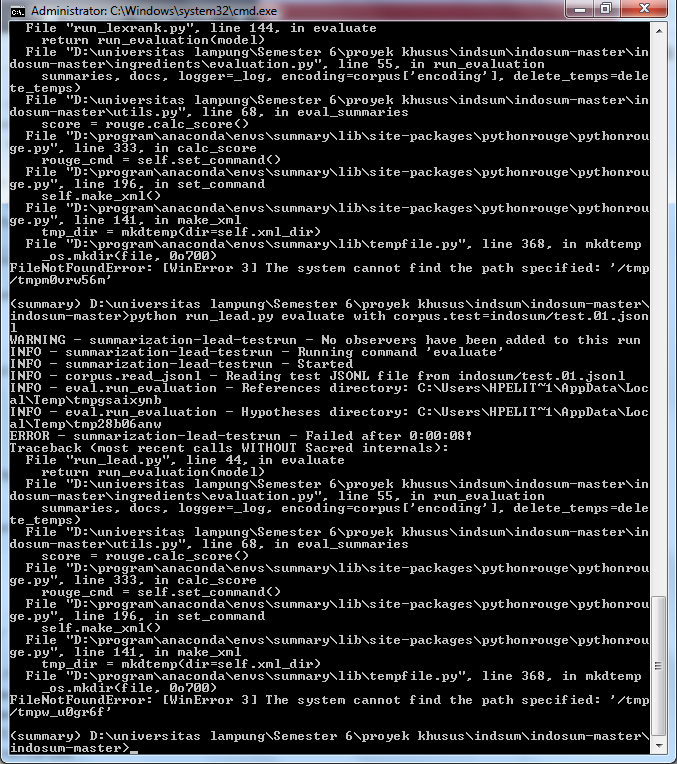






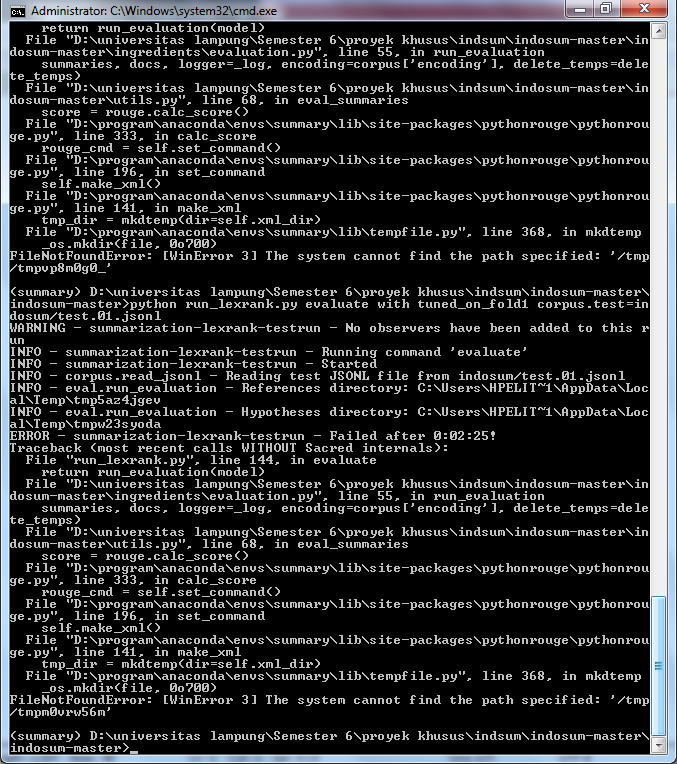
Dari hasil runingan data ini dapat dilihat bahwa menghasilkan suatu nilai gausian elimination yang dapat dilihat nilainya dan setiap tabel ada nilai, dan ada yang tidak. Terdapat 26 state, dan terdapat nilai rata-rata dari dictionary yang artinya rata-rata suatu kata, dan memiliki nilai gama. Jadi fungsi dari sourcode ini melihat nilai gausian dari kata dan rata-ratanya. Dan memberitahukan 26 state dalam prosesnya.

**Gambar 10. Fungsi dari run\_lead.py**



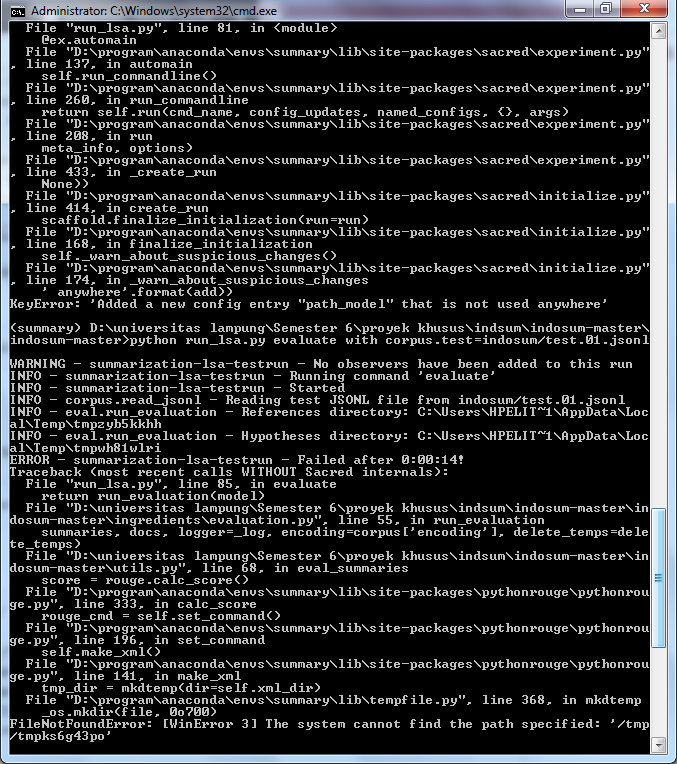
Hasilnya tidak ada, tidak bisa menggunakan fungsi yang sama dalam meruning program dan terdapat enviromentnya tidak suport.

**Gambar 11. Fungsi dari run\_lexrank.py**



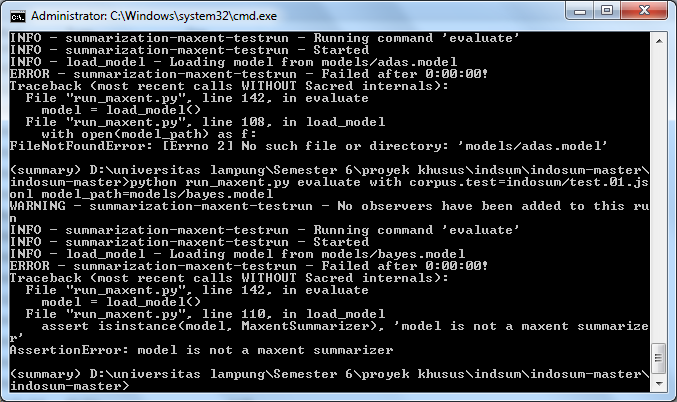
Dalam sourcode ini terdapat eror ketika dijalankan, terdapat suatu baris yang eror didalam library enviromentnya yang tidak suport.

**Gambar 12. Fungsi dari run\_lsa.py**



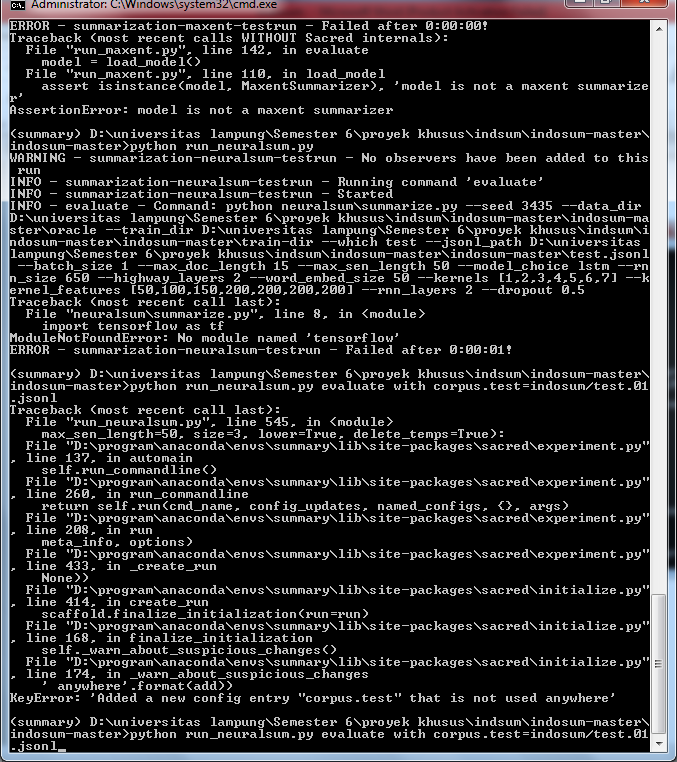
Terdapat eror ketika dijalankan pada data tersebut sourcode tersebut.

**Gambar 13. Fungsi dari run\_maxent.py**



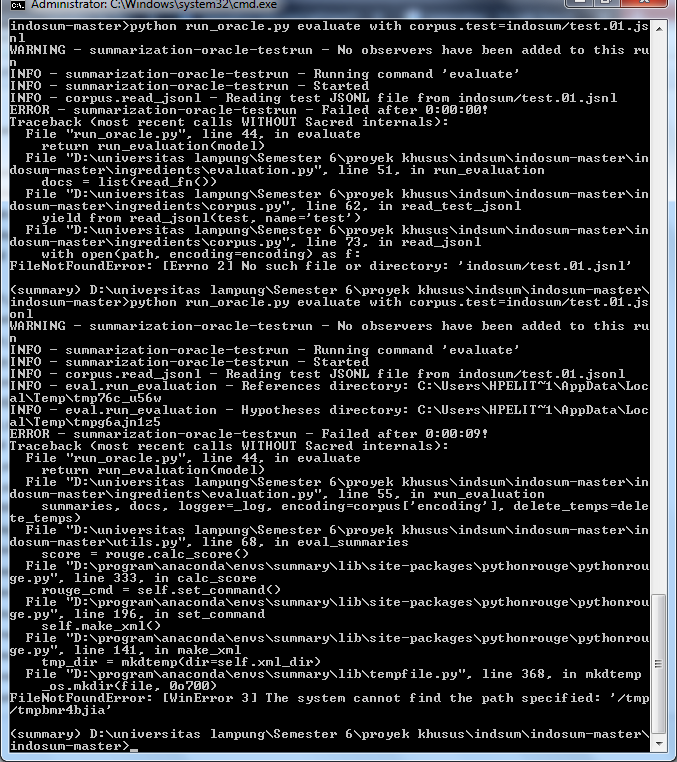
Tidak bisa dijalankan terjadi eror dalam meruning programnya.

**Gambar 14. Fungsi dari run\_neuralsum.py**



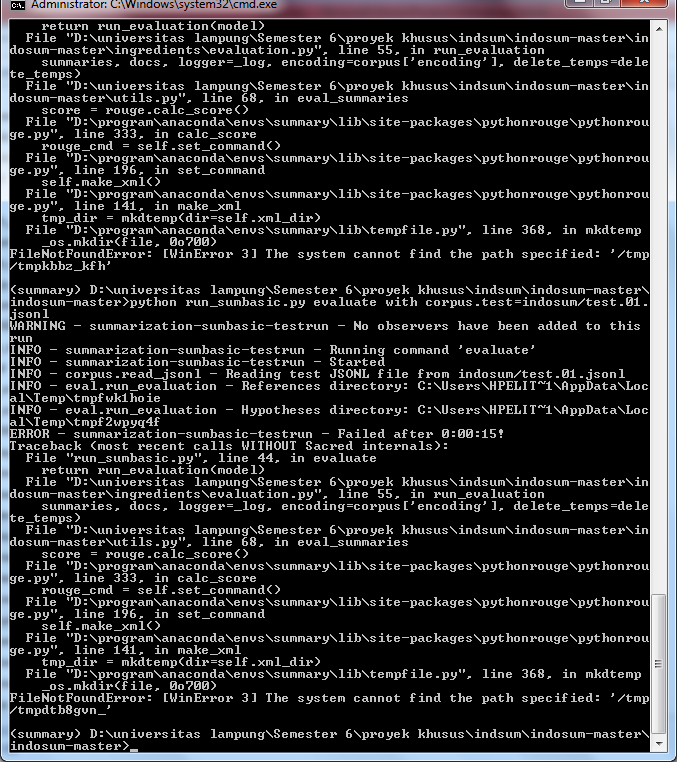
Fungsinya eror dalam menjalankannya.

**Gambar 15. Fungsi dari run\_oracle.py**



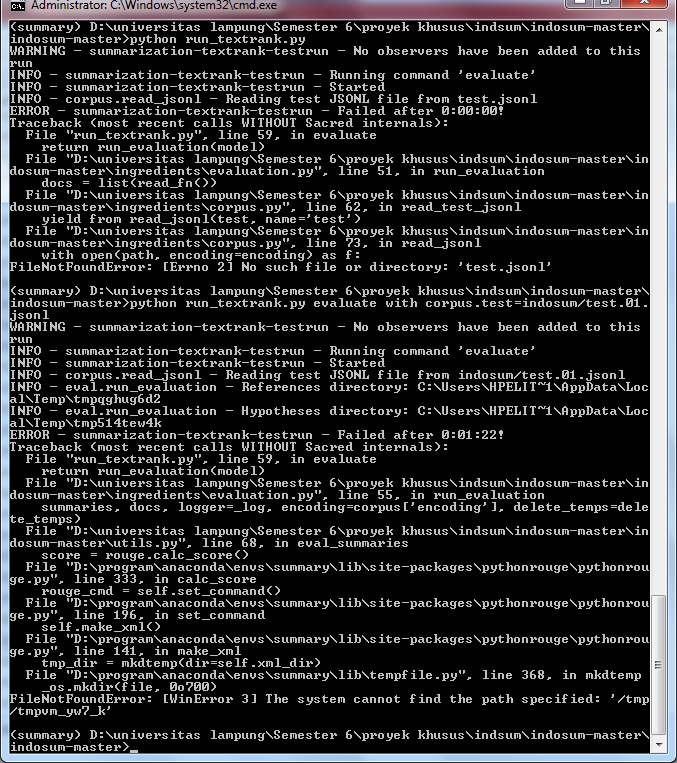
Dalam menjalan sourcodenya fungsi tersebut eror.

**Gambar 16. Fungsi dari run\_sumbasic.py**



Terjadi eror saat menjalankan program dikarenakan terdapat path yang tidak support dalam environment tersebut.

**Gambar 17. Fungsi dari run\_textrank.py**



Terjadi eror yang sama saat menjalankan program, dikarenakan environment terdapat masalah eror.