

Οδηγίες Εκτέλεσης

Το πρόγραμμα δέχεται ως είσοδο:

1. Το σύνολο των αρχείων.
2. Το σύνολο των θεματικών ομάδων.
3. Το αριθμό των επαναλήψεων.
4. Ένα path εισόδου.
5. Ένα path εξόδου.

Στο path εισόδου πρέπει να περιέχει μέσα σε ένα φάκελο input να περιέχει:

- Το αρχείο κειμένου με τα κέντρα με όνομα centers.
- Το αρχείο stopwords.
- Ένα φάκελο docs που περιέχει τα κείμενα με ονόματα (0 έως N-1)

Τα αρχεία κειμένου πρέπει να μην έχουν κατάληξη .txt στο όνομα τους.

Το αρχείο centers πρέπει να έχει την μορφή:

<ομάδα> <διανυσμα>

π.χ. 3 0 1 0 1 0

Παράδειγμα εκτέλεσης:

```
hadoop jar /home/kds.jar 5 5 10 /user/hduser/project/input  
/user/hduser/project/output
```