

```

// Input file for word count
input_file = sc.textFile("/home/ubuntu/server/Uploads/tweet_count.txt")
input_file = sc.textFile("/home/ubuntu/server/Uploads/news_count.txt")

//List words for frequency count
words = ["education","canada","university","dalhousie","expensive","good school","good schools","bad schools","bad school","poor school","poor schools","faculty","computer","science","graduate"]

broadcast_words = sc.broadcast(words)

// split the lines of file using space
map = input_file.flatMap(lambda line: line.split(" "))

//filter list of words from main input file
m = map.filter(lambda x: x in broadcast_words .value)

//Get count of filtered word
m1 = m.map(lambda word: (word, 1))

// Merge count of the same words
counts = m1.reduceByKey(lambda a, b: a + b)

// Save the result in file
counts.saveAsTextFile("/home/ubuntu/server/Uploads/output3")

```

Refernces

[1]"Python Spark Shell - PySpark - Word Count Example", TutorialKart, 2020. [Online]. Available: <https://www.tutorialkart.com/apache-spark/python-spark-shell-pyspark-example/>. [Accessed: 27- Mar- 2020].