

KDM Assignment -1

Student id:16269904

Class id:4

Name:Lakshmi Korrapati

Task 1

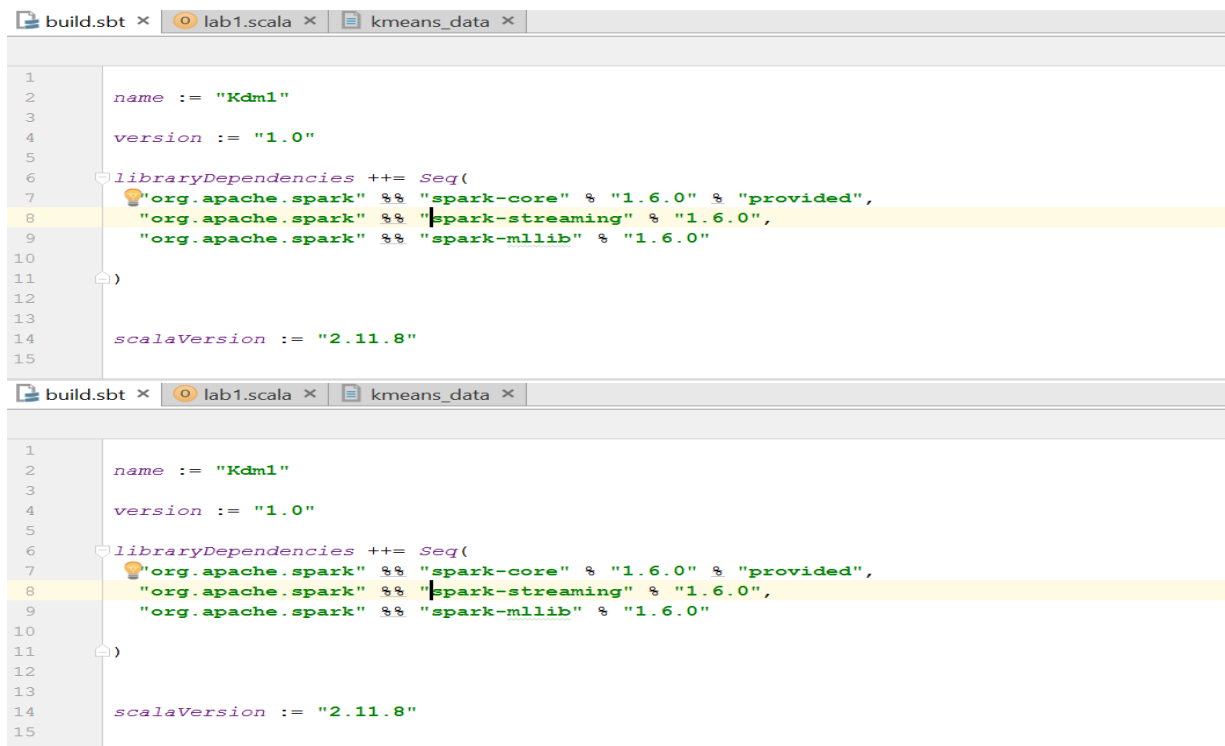
Create a Scala project In IntelliJ using Scala and SBT. Add any library to SBT and use in it in your project.

What is SBT?

SBT is an open source build tool for Scala and Java projects, similar to Java's Maven or Ant. Its main features are: native support for compiling Scala code and integrating with many Scala test frameworks.

Adding Library to SBT

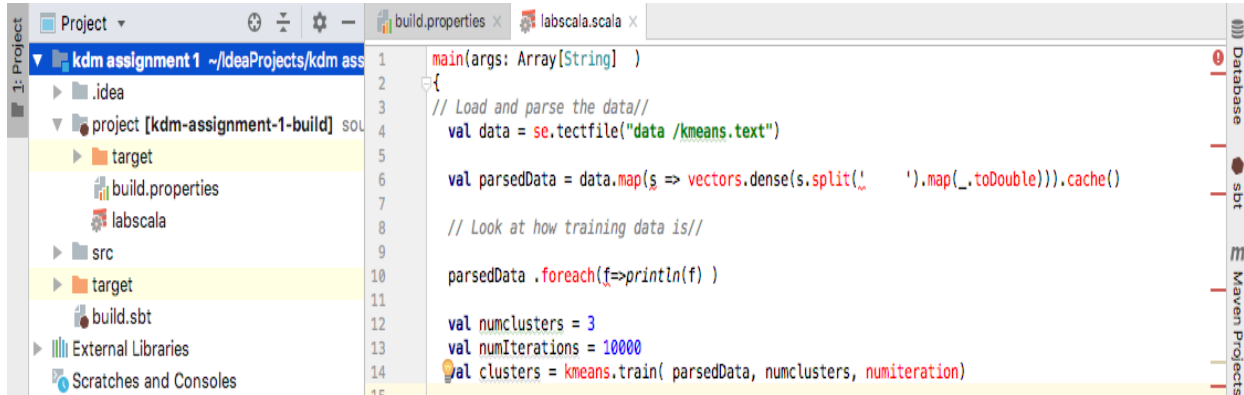
In the build.sbt I have added the machine learning libraries(mllib) and I have used them to write a simple kmeans clustering program which can be used in NLP for grouping similar words.



```
1  name := "Kdm1"
2
3
4  version := "1.0"
5
6  libraryDependencies += Seq(
7    "org.apache.spark" %% "spark-core" % "1.6.0" % "provided",
8    "org.apache.spark" %% "spark-streaming" % "1.6.0",
9    "org.apache.spark" %% "spark-mllib" % "1.6.0"
10 )
11
12
13
14  scalaVersion := "2.11.8"
15
```

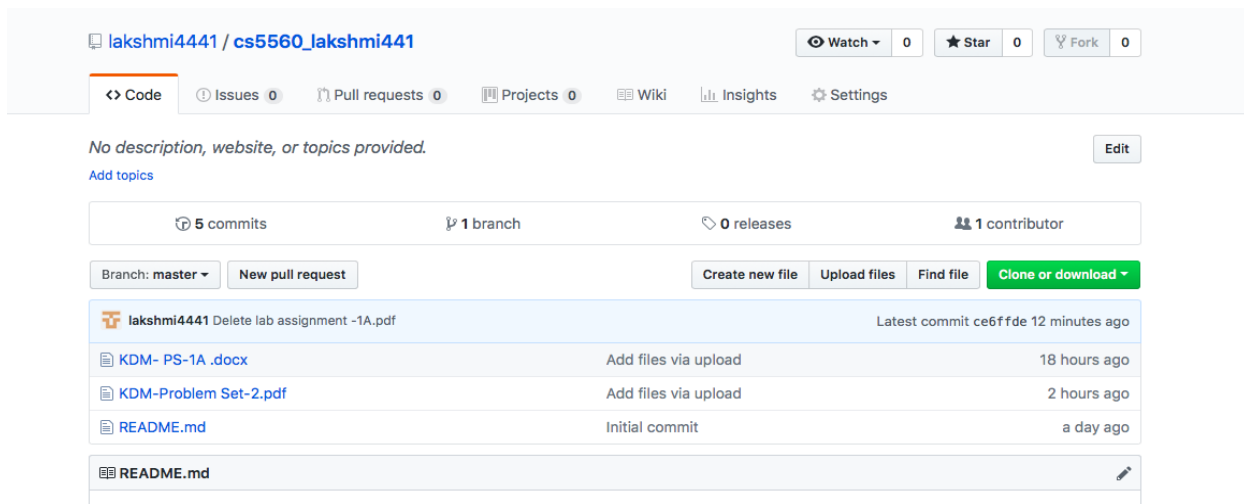
Scala Kmeans Clustering Using mllib

The below kmeans clustering program group similar elements which is used in Bag Of Words or grouping similar text during Q & A.



```
1 main(args: Array[String]) {
2   // Load and parse the data//
3   val data = se.textfile("data /kmeans.text")
4
5   val parsedData = data.map(s => vectors.dense(s.split(' ').map(_.toDouble))).cache()
6
7   // Look at how training data is//
8
9   parsedData.foreach(f=>println(f))
10
11
12   val numclusters = 3
13   val numIterations = 10000
14   val clusters = kmeans.train(parsedData, numclusters, numiteration)
```

2. create Git Hub account and wiki page



lakshmi4441 / cs5560_lakshmi441

Watch 0 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

No description, website, or topics provided. [Edit](#)

[Add topics](#)

5 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Create new file Upload files Find file Clone or download

File	Commit	Time
KDM- PS-1A .docx	ce6ffde	18 hours ago
KDM-Problem Set-2.pdf	ce6ffde	2 hours ago
README.md	Initial commit	a day ago

README.md

 [lakshmi4441](#) / [cs5560_lakshmi441](#)

 Watch 0  Star 0  Fork 0

[Code](#) [Issues 0](#) [Pull requests 0](#) [Projects 0](#) [Wiki](#) [Insights](#) [Settings](#)

Create new page

Home


Write


Preview


h1


h2


h3


























Edit mode:

Markdown

Welcome to the cs5560_lakshmi441 wiki!