

Practical No. 4

Aim: Write a Spark code for the given application and handle error

Theory: Scala offers different classes for functional error handling in Spark. These classes include but are not limited to Try/Success/Failure, Option/Some/None, Either/Left/Right. Depending on what you are trying to achieve you may want to choose a trio class based on the unique expected outcome of your code.

Implementation:

switch to Hadoop user

Install scala-sbt via the latest command available in the terminal

```
echo "deb https://repo.scala-sbt.org/scalasbt/debian all main" | sudo tee /etc/apt/sources.list.d/sbt.list
echo "deb https://repo.scala-sbt.org/scalasbt/debian /" | sudo tee /etc/apt/sources.list.d/sbt_old.list
curl -sL "https://keyserver.ubuntu.com/pks/lookup?op=get&search=0x2EE0EA64E40A89B84B2DF73499E82A75642AC823" | sudo apt-key add
sudo apt-get update
sudo apt-get install sbt
```

enter sbt in the terminal to verify installation

create the scala program for exception handling

\$ nano ExceptionHandlingTest.scala

```
import org.apache.spark.sql.SparkSession

object ExceptionHandlingTest {
  def main(args: Array[String]): Unit = {
    val spark = SparkSession
      .builder
      .appName("ExceptionHandlingTest")
      .getOrCreate()

    spark.sparkContext.parallelize(0 until
      spark.sparkContext.defaultParallelism).foreach { i =>
```

```

        if (math.random > 0.75) {
            throw new Exception("Testing exception handling")
        }
    }

    spark.stop()
}
}

```

create the sbt dependency file 'exceptionhandlingtest.sbt'

```
name := "exceptionhandlingtest"
```

```
version := "1.0"
```

```
scalaVersion := "2.12.15"
```

```
val sparkVersion = "3.3.1"
```

```
libraryDependencies ++= Seq(
    "org.apache.spark" %% "spark-core" % sparkVersion,
    "org.apache.spark" %% "spark-sql" % sparkVersion
)
```

start spark master & worker

```
$ spark-master.sh
```

```
$ spark-worker.sh spark://Ubuntu.myguest.virtualbox.org:7077/
```

go back to the terminal pointing to the directory with scala and sbt file and create the sbt package

```
$ sbt package
```

submit the program to spark

```
$ spark-submit --class 'ExceptionHandlingTest' --master
'spark://Ubuntu.myguest.virtualbox.org:7077/' \target/scala-2.12/exceptionhandling-2.12-1.0.jar
```