Troy Yang

DS4300 Final Project Report

4/19/2018

Following your feedback, I decided to really focus in on edge detection in terms of where I should be focusing my efforts for this project. I had implemented 2 custom transformers, a raw pixel featurizer which took the path of an image and convert it to a vector and labeler that took the path of an image and returns a label (1 for dog 0 for cat). In addition to those transformers, I added three more for the final submission. Each does some sort of edge detection/filtering. Each featurizer applies the edge detection algorithm mentioned in their name, so the Canny edge featurizer uses the Canny edge detection algorithm, Sobel edge featurizer uses a Sobel filter (applying a Sobel filter is actually one of the steps in the Canny edge detection algorithm), and Laplace edge featurizer uses a Laplacian filter. In each of those transformers, the altered image is then turned into a vector. I then created different pipelines which consisted of different combinations of the featurizers and classifiers. The classifiers I used were all ones that came for free in MLlib. I used Naive Bayes, Linear Support Vector Machines and a Multilayer Perceptron Classifier as these machine learning algorithms tend to do well with data that has high dimensionality. Unfortunately, the different pipelines all yielded similar accuracy in the range of 55-60%.

Each transformer was its own class, and then I had a main script where I actually created and ran different pipelines. In my main I would run one pipeline, and then I would I would just swap different parts of the pipeline to try different combinations. I initially tried running one pipeline after another, but since one pipeline already took a long time to run, I decided against that.

Bytedeco is an ongoing project that makes C/C++ libraries available to the Java platform. They have a library called JavaCV which is a wrapper for the OpenCV library (computer vision library). This was incredibly useful for me when it came to creating the edge detection transformers. While there was some useful documentation and examples online, there (understandably) weren't a ton of Stack Overflow questions, so when I ran into problems sometimes things got a little dicey. The website aishack.in was also extremely useful for me, as it was able to explain and help me understand what exactly a lot of the edge detection algorithms did (especially compared to some other websites I looked at).

Despite lackluster results, I am still satisfied with the outcomes of this project. I definitely learned a lot, whether it was about using libraries such as MLlib or JavaCV and integrating them, or the different edge detection algorithms I used in transformers. Another part of the projected I really enjoyed (definitely not in the moment) was being forced to debug certain things, and not have a readily available answer on stack overflow. Obviously this lead to a slower pace of work, but this made me feel more confident in my ability to "figure stuff out". There were several while working on this project where something wouldn't work and I thought to myself that there was no way I was going to get something to work. Being able to push through problems like that and figuring out what was wrong was immensely gratifying, even if sometimes the problems were pretty stupid, like in one case where I was drawing data onto the wrong image.

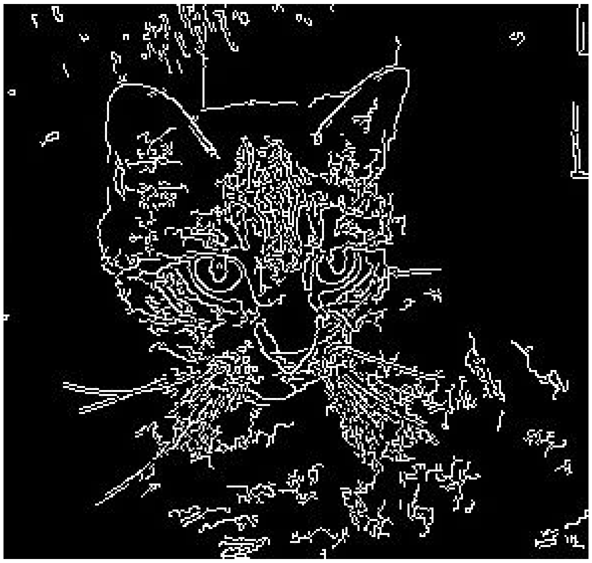
Exploring different technologies and learning about image processing was also something that I underestimated in terms of how long it would take me to do, but was obviously time well spent.

I've attached the and labeled the outputs of some of the pipelines I ran (I didn't include all outputs because I forgot to take screenshots for some outputs, and it takes awhile for each pipeline to run, but I can go back and rerun them and send you those screenshots as well if you want me to), also well as examples of the edge detection algorithms used.

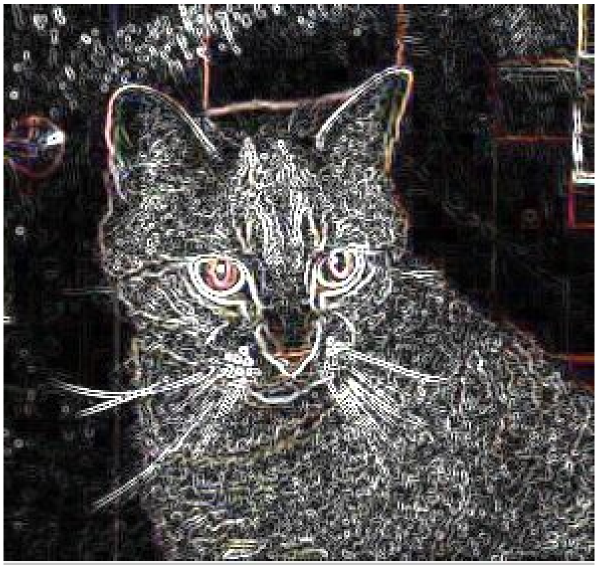
Image Transformations (Edges)

Original

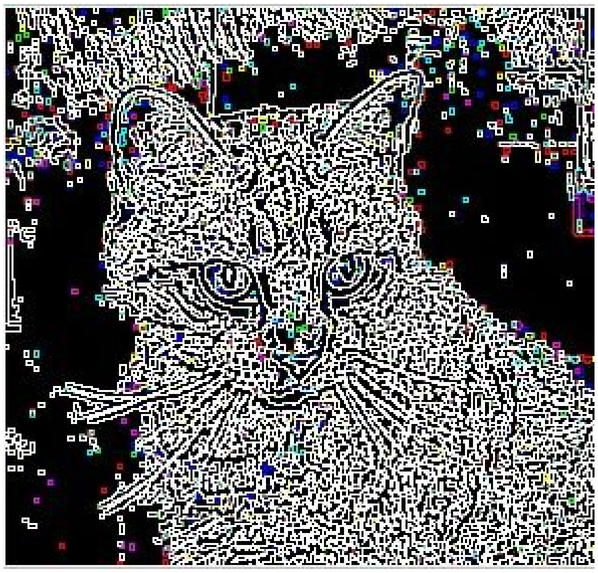
Canny



Sobel

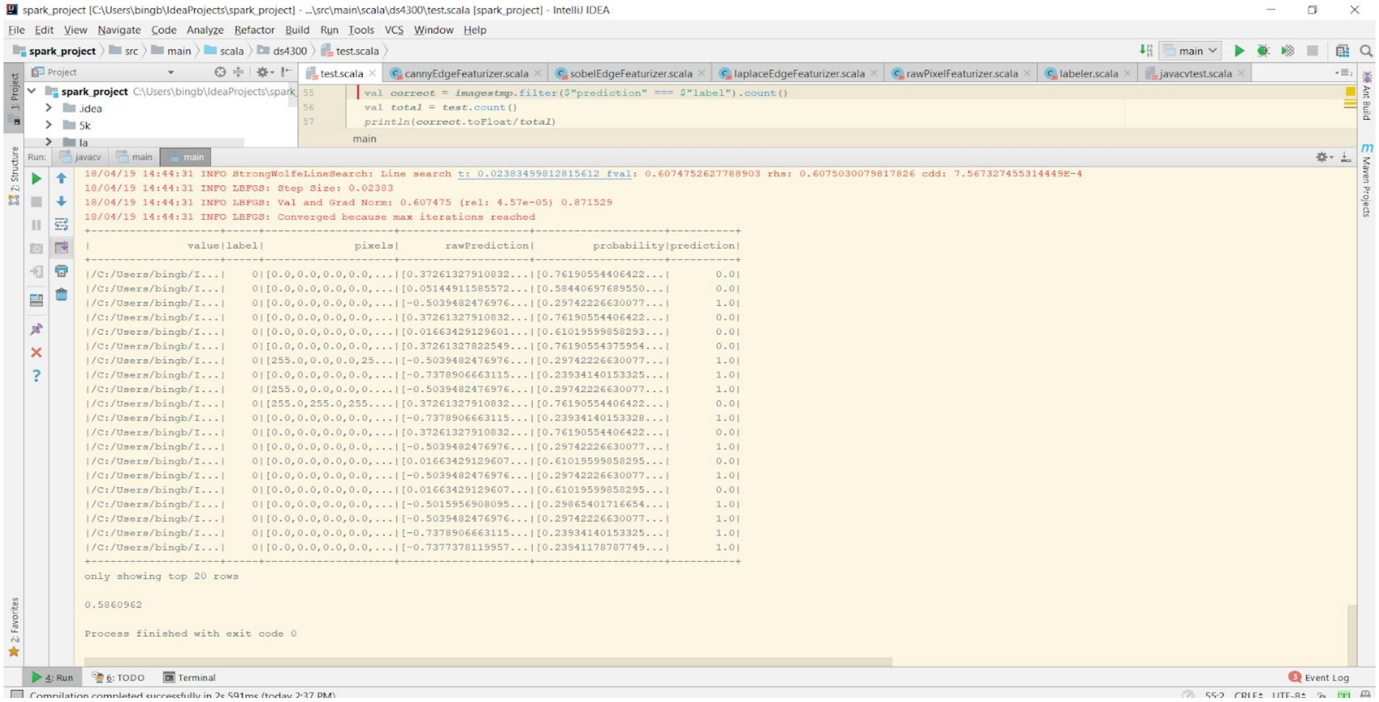


Laplace



Outputs

Canny + MPC

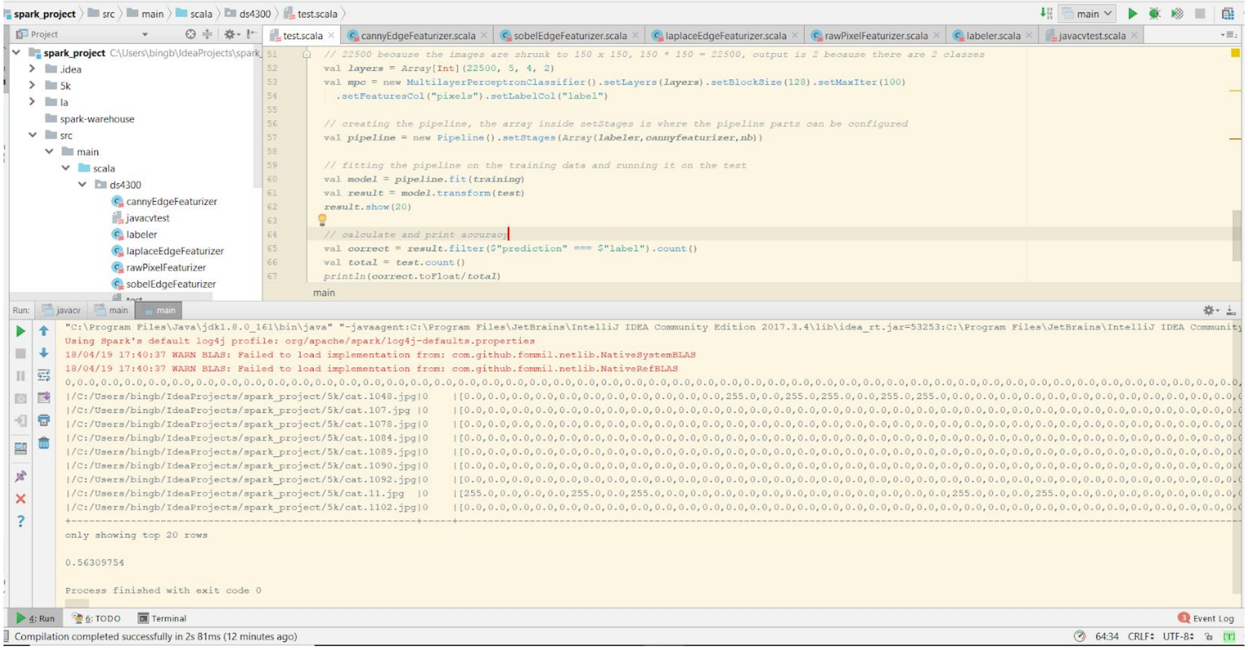

spark_project [C:\Users\bingb\IdeaProjects\spark_project] - ...\src\main\scala\ds4300\test.scala [spark_project] - IntelliJ IDEA
File Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
spark_project src main scala ds4300 test.scala
Project
spark_project C:\Users\bingb\IdeaProjects\spark_
.idea
5k
la
test.scala
cannyEdgeFeaturizer.scala
sobelEdgeFeaturizer.scala
laplaceEdgeFeaturizer.scala
rawPixelFeaturizer.scala
labeler.scala
javacvtest.scala
main
Run: javacv main main
4: Run
6: TODO
Terminal
Event Log

```
val correct = imagestmp.filter($"prediction" === $"label").count()
val total = test.count()
println(correct.toFloat/total)
```

```
18/04/19 14:44:31 INFO StrongWolfeLineSearch: Line search t: 0.02383499812815612 fval: 0.6074752627788903 rhs: 0.6075030079817826 cdd: 7.567327455314449E-4
18/04/19 14:44:31 INFO LBFGS: Step Size: 0.02383
18/04/19 14:44:31 INFO LBFGS: Val and Grad Norm: 0.607475 (rel: 4.57e-05) 0.871529
18/04/19 14:44:31 INFO LBFGS: Converged because max iterations reached
+--------------------+-----+--------------------+--------------------+--------------------+----------+
|               value|label|              pixels|       rawPrediction|         probability|prediction|
+--------------------+-----+--------------------+--------------------+--------------------+----------+
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.37261327910832...|[0.76190554406422...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.05144911585572...|[0.58440697689550...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.37261327910832...|[0.76190554406422...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.01663429129601...|[0.61019599858293...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.37261327822549...|[0.76190554375954...|       0.0|
|/C:/Users/bingb/I...|    0|[255.0,0.0,0.0,25...|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.7378906663115...|[0.23934140153325...|       1.0|
|/C:/Users/bingb/I...|    0|[255.0,0.0,0.0,0....|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[255.0,255.0,255....|[0.37261327910832...|[0.76190554406422...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.7378906663115...|[0.23934140153328...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.37261327910832...|[0.76190554406422...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.01663429129607...|[0.61019599858295...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.01663429129607...|[0.61019599858295...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5015956908095...|[0.29865401716654...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5039482476976...|[0.29742226630077...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.7378906663115...|[0.23934140153325...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.7377378119957...|[0.23941178787749...|       1.0|
+--------------------+-----+--------------------+--------------------+--------------------+----------+
only showing top 20 rows

0.5860962

Process finished with exit code 0
```

Compilation completed successfully in 2s 591ms (today 2:37 PM)

Canny + NB



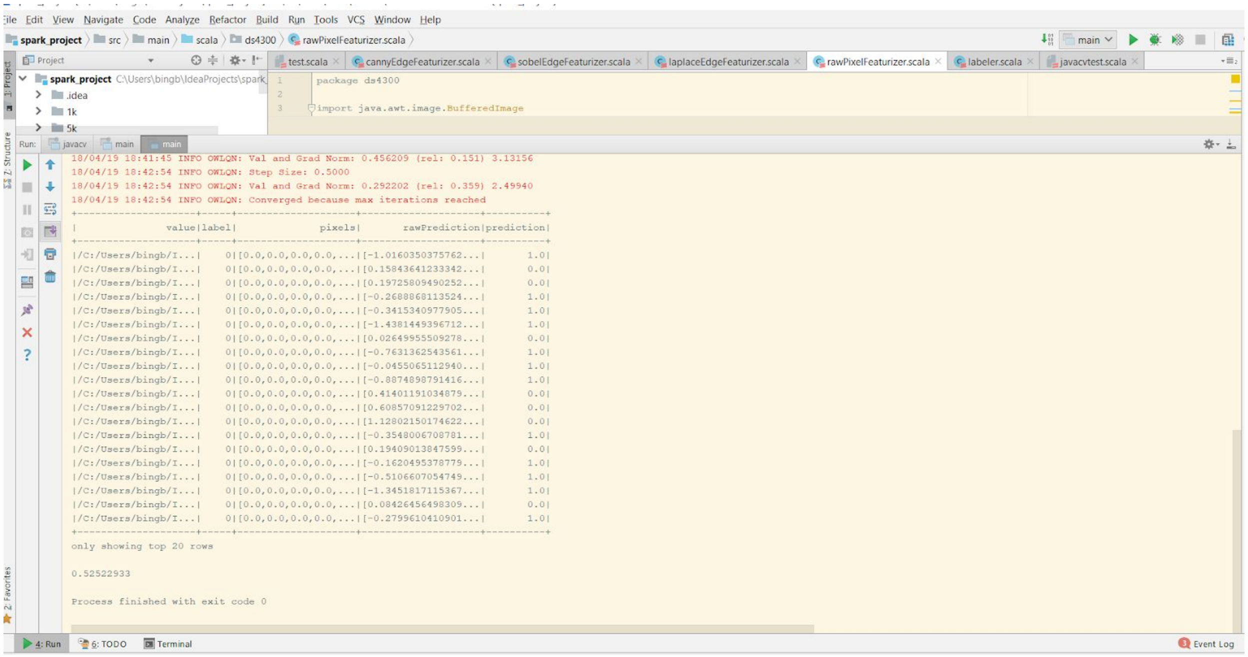
```scala
// 22500 because the images are shrunk to 150 x 150, 150 * 150 = 22500, output is 2 because there are 2 classes
val layers = Array[Int](22500, 5, 4, 2)
val mpc = new MultilayerPerceptronClassifier().setLayers(layers).setBlockSize(128).setMaxIter(100)
  .setFeaturesCol("pixels").setLabelCol("label")

// creating the pipeline, the array inside setStages is where the pipeline parts can be configured
val pipeline = new Pipeline().setStages(Array(labeler,cannyfeaturizer,nb))

// fitting the pipeline on the training data and running it on the test
val model = pipeline.fit(training)
val result = model.transform(test)
result.show(20)

// calculate and print accuracy
val correct = result.filter($"prediction" === $"label").count()
val total = test.count()
println(correct.toFloat/total)
```

main

```
"C:\Program Files\Java\jdk1.8.0_161\bin\java" "-javaagent:C:\Program Files\JetBrains\IntelliJ IDEA Community Edition 2017.3.4\lib\idea_rt.jar=53253:C:\Program Files\JetBrains\IntelliJ IDEA Community
Using Spark's default log4j profile: org/apache/spark/log4j-defaults.properties
18/04/19 17:40:37 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.NativeSystemBLAS
18/04/19 17:40:37 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.NativeRefBLAS
0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1048.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,255.0,0.0,255.0,255.0,0.0,255.0,255.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.107.jpg |0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1078.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1084.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1089.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1090.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1092.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.11.jpg  |0    |[255.0,0.0,0.0,0.0,0.0,255.0,0.0,0.0,255.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,255.0,0.0,0.0,0.0,0.0,255.0,0.0,0.0,0.0,0.0,0.0,0.0,0
|/C:/Users/bingb/IdeaProjects/spark_project/5k/cat.1102.jpg|0    |[0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0.0,0
+-----------------------------------------------------------+-----+----------------------------------------------------------------------------------------------------------------------------------
only showing top 20 rows

0.56309754

Process finished with exit code 0
```

Compilation completed successfully in 2s 81ms (12 minutes ago)

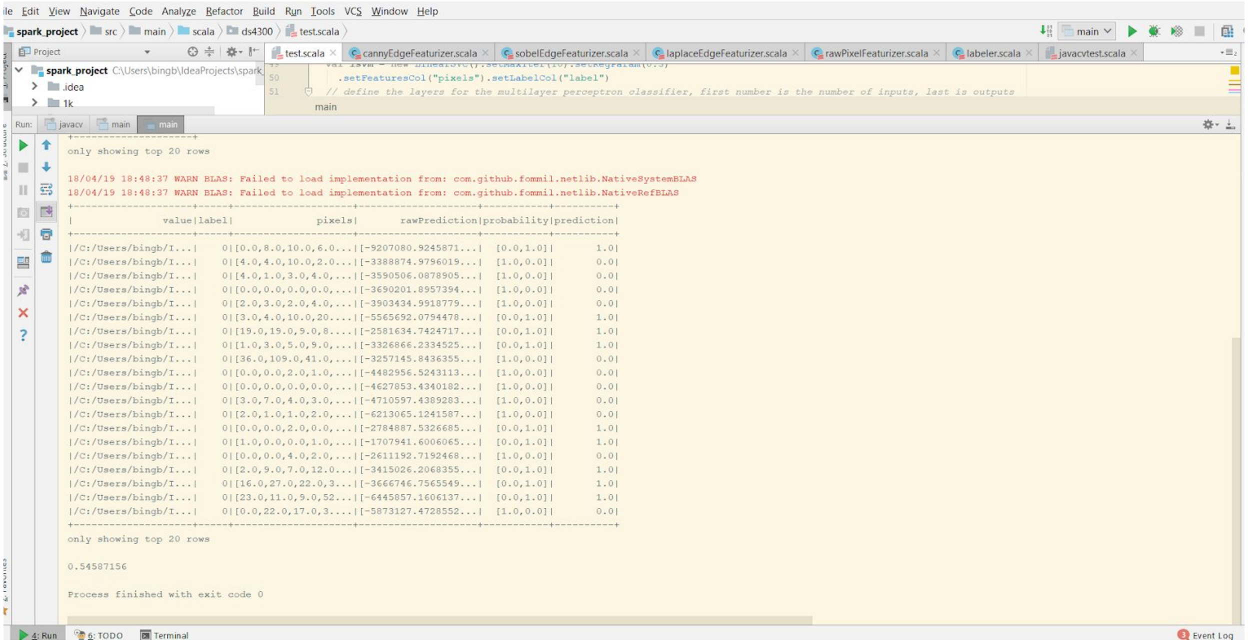
Canny + LSVM



```scala
package ds4300

import java.awt.image.BufferedImage
```

```
18/04/19 18:41:45 INFO OWLQN: Val and Grad Norm: 0.456209 (rel: 0.151) 3.13156
18/04/19 18:42:54 INFO OWLQN: Step Size: 0.5000
18/04/19 18:42:54 INFO OWLQN: Val and Grad Norm: 0.292202 (rel: 0.359) 2.49940
18/04/19 18:42:54 INFO OWLQN: Converged because max iterations reached
+--------------------+-----+--------------------+--------------------+----------+
|               value|label|              pixels|       rawPrediction|prediction|
+--------------------+-----+--------------------+--------------------+----------+
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.0160350375762...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.15843641233342...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.19725809490252...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.2688868113524...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.3415340977905...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.4381449396712...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.02649955509278...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.7631362543561...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.0455065112940...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.8874898791416...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.41401191034879...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.60857091229702...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[1.12802150174622...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.3548006708781...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.19409013847599...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.1620495378779...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.5106607054749...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.3451817115367...|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[0.08426456498309...|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-0.2799610410901...|       1.0|
+--------------------+-----+--------------------+--------------------+----------+
only showing top 20 rows

0.52522933

Process finished with exit code 0
```

Sobel + NB

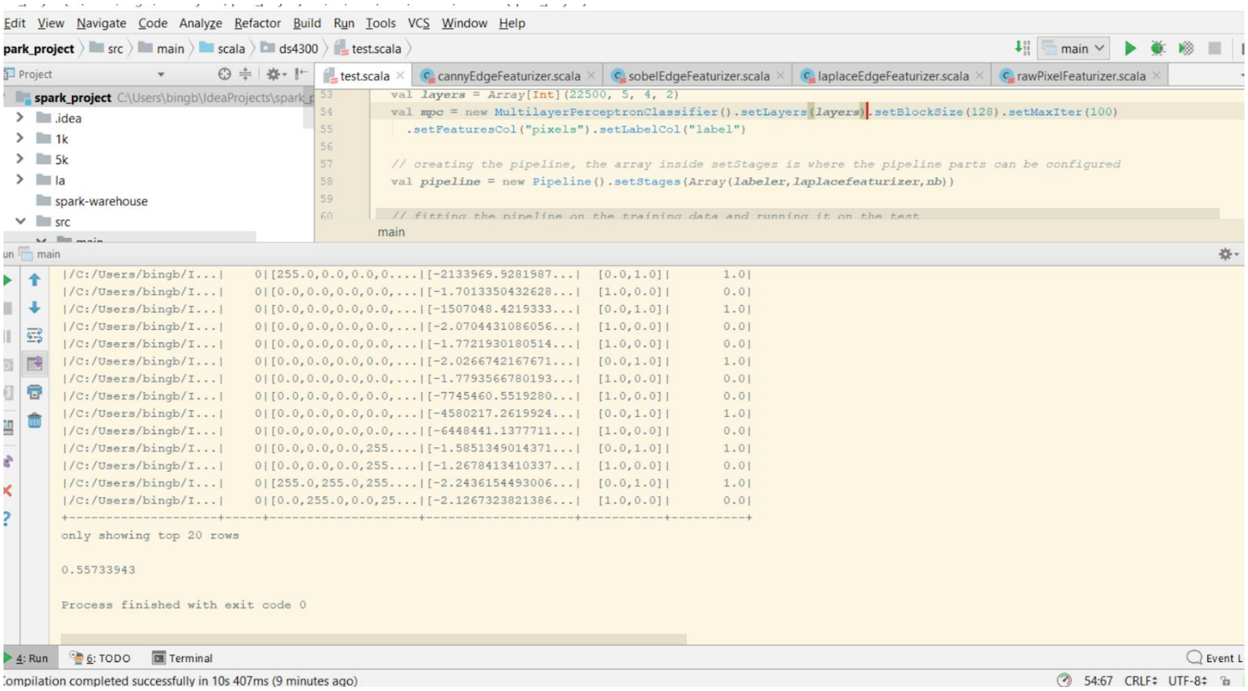


ile Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
spark_project › src › main › scala › ds4300 › test.scala
main
Project
spark_project C:\Users\bingb\IdeaProjects\spark_
.idea
1k
test.scala
cannyEdgeFeaturizer.scala
sobelEdgeFeaturizer.scala
laplaceEdgeFeaturizer.scala
rawPixelFeaturizer.scala
labeler.scala
javacvtest.scala
main
Run:
javacv
main
main
4: Run
6: TODO
Terminal
Event Log

```
    val lsvm = new LinearSVC().setMaxIter(10).setRegParam(0.3)
      .setFeaturesCol("pixels").setLabelCol("label")
    // define the layers for the multilayer perceptron classifier, first number is the number of inputs, last is outputs
```

```
+--------------------+
only showing top 20 rows

18/04/19 18:48:37 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.NativeSystemBLAS
18/04/19 18:48:37 WARN BLAS: Failed to load implementation from: com.github.fommil.netlib.NativeRefBLAS
+--------------------+-----+--------------------+--------------------+-----------+----------+
|               value|label|              pixels|       rawPrediction|probability|prediction|
+--------------------+-----+--------------------+--------------------+-----------+----------+
|/C:/Users/bingb/I...|    0|[0.0,8.0,10.0,6.0...|[-9207080.9245871...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[4.0,4.0,10.0,2.0...|[-3388874.9796019...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[4.0,1.0,3.0,4.0,...|[-3590506.0878905...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-3690201.8957394...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[2.0,3.0,2.0,4.0,...|[-3903434.9918779...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[3.0,4.0,10.0,20....|[-5565692.0794478...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[19.0,19.0,9.0,8....|[-2581634.7424717...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[1.0,3.0,5.0,9.0,...|[-3326866.2334525...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[36.0,109.0,41.0,...|[-3257145.8436355...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,2.0,1.0,...|[-4482956.5243113...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-4627853.4340182...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[3.0,7.0,4.0,3.0,...|[-4710597.4389283...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[2.0,1.0,1.0,2.0,...|[-6213065.1241587...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,2.0,0.0,...|[-2784887.5326685...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[1.0,0.0,0.0,1.0,...|[-1707941.6006065...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,4.0,2.0,...|[-2611192.7192468...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[2.0,9.0,7.0,12.0...|[-3415026.2068355...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[16.0,27.0,22.0,3...|[-3666746.7565549...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[23.0,11.0,9.0,52...|[-6445857.1606137...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,22.0,17.0,3....|[-5873127.4728552...|  [1.0,0.0]|       0.0|
+--------------------+-----+--------------------+--------------------+-----------+----------+
only showing top 20 rows

0.54587156

Process finished with exit code 0
```

Laplace + NB

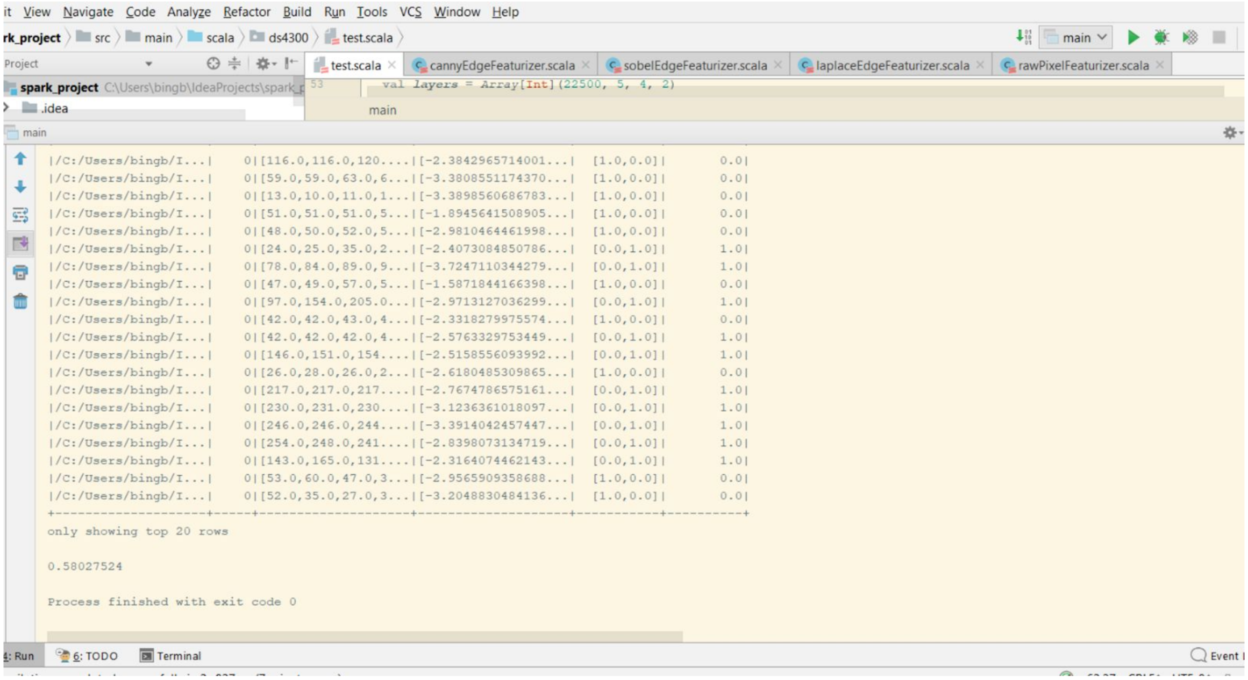

Edit View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
park_project src main scala ds4300 test.scala
main
Project
test.scala
cannyEdgeFeaturizer.scala
sobelEdgeFeaturizer.scala
laplaceEdgeFeaturizer.scala
rawPixelFeaturizer.scala
spark_project C:\Users\bingb\IdeaProjects\spark_p
.idea
1k
5k
la
spark-warehouse
src
main
4: Run
6: TODO
Terminal
Event L
54:67 CRLF UTF-8

```
    val layers = Array[Int](22500, 5, 4, 2)
    val mpc = new MultilayerPerceptronClassifier().setLayers(layers).setBlockSize(128).setMaxIter(100)
      .setFeaturesCol("pixels").setLabelCol("label")

    // creating the pipeline, the array inside setStages is where the pipeline parts can be configured
    val pipeline = new Pipeline().setStages(Array(labeler,laplacefeaturizer,nb))

    // fitting the pipeline on the training data and running it on the test
```

```
|/C:/Users/bingb/I...|    0|[255.0,0.0,0.0,0....|[-2133969.9281987...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.7013350432628...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1507048.4219333...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-2.0704431086056...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.7721930180514...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-2.0266742167671...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-1.7793566780193...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-7745460.5519280...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-4580217.2619924...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,0.0,...|[-6448441.1377711...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,255....|[-1.5851349014371...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,0.0,0.0,255....|[-1.2678413410337...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[255.0,255.0,255....|[-2.2436154493006...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[0.0,255.0,0.0,25...|[-2.1267323821386...|  [1.0,0.0]|       0.0|
+--------------------+-----+--------------------+--------------------+----------+----------+
only showing top 20 rows

0.55733943

Process finished with exit code 0
```

Compilation completed successfully in 10s 407ms (9 minutes ago)

Raw + NB


it View Navigate Code Analyze Refactor Build Run Tools VCS Window Help
rk_project src main scala ds4300 test.scala
main
Project
test.scala
cannyEdgeFeaturizer.scala
sobelEdgeFeaturizer.scala
laplaceEdgeFeaturizer.scala
rawPixelFeaturizer.scala
spark_project C:\Users\bingb\IdeaProjects\spark_p
.idea
main
4: Run
6: TODO
Terminal
Event

```
    val layers = Array[Int](22500, 5, 4, 2)
```

```
|/C:/Users/bingb/I...|    0|[116.0,116.0,120....|[-2.3842965714001...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[59.0,59.0,63.0,6...|[-3.3808551174370...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[13.0,10.0,11.0,1...|[-3.3898560686783...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[51.0,51.0,51.0,5...|[-1.8945641508905...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[48.0,50.0,52.0,5...|[-2.9810464461998...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[24.0,25.0,35.0,2...|[-2.4073084850786...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[78.0,84.0,89.0,9...|[-3.7247110344279...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[47.0,49.0,57.0,5...|[-1.5871844166398...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[97.0,154.0,205.0...|[-2.9713127036299...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[42.0,42.0,43.0,4...|[-2.3318279975574...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[42.0,42.0,42.0,4...|[-2.5763329753449...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[146.0,151.0,154....|[-2.5158556093992...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[26.0,28.0,26.0,2...|[-2.6180485309865...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[217.0,217.0,217....|[-2.7674786575161...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[230.0,231.0,230....|[-3.1236361018097...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[246.0,246.0,244....|[-3.3914042457447...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[254.0,248.0,241....|[-2.8398073134719...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[143.0,165.0,131....|[-2.3164074462143...|  [0.0,1.0]|       1.0|
|/C:/Users/bingb/I...|    0|[53.0,60.0,47.0,3...|[-2.9565909358688...|  [1.0,0.0]|       0.0|
|/C:/Users/bingb/I...|    0|[52.0,35.0,27.0,3...|[-3.2048830484136...|  [1.0,0.0]|       0.0|
+--------------------+-----+--------------------+--------------------+----------+----------+
only showing top 20 rows

0.58027524

Process finished with exit code 0
```