

UNIVERSITY OF HASSELT

BACHELOR THESIS: EXPERIMENTS

Machine learning techniques for flow-based network intrusion detection systems: Experiments

Author:
Axel FAES

Advisor:
Prof. Dr. Peter QUAX
Prof. Dr. Wim LAMOTTE
Mentor:
Bram BONNE
Pieter ROBYNS

Academiejaar 2015-2016
June, 2016

universiteit
▶▶ hasselt

KNOWLEDGE IN ACTION

Contents

1	Introduction	1
2	Baseline	3
3	K-Nearest Neighbors	5
4	Decision Tree Classifier	7
4.1	Standard feature set	7
4.2	TCP feature set	14
4.3	Country feature set	20
5	Naive Bayes	27
6	Support Vector Machines with Linear Kernel	29
7	Support Vector Machine with RBF Kernel	31
8	Neural Networks	33
9	One class Support Vector Machines	35

Chapter 1

Introduction

This document contains all the different results from the experiments that have been done for the thesis. Each chapter contains the results from the algorithm that is being described. Each algorithm has been tested using **3 different feature sets**. Each experiment has been repeated **10 times**.

Chapter 2

Baseline

Chapter 3

K-Nearest Neighbors

Chapter 4

Decision Tree Classifier

4.1 Standard feature set

4.1.1 CTU Dataset

Table 4.1: DecisionTreeClassifier with Standard feature set CTU:
Experiment 1.

Multi-class F-score	0.600550604122
Multi-class Precision	0.657768055449
Multi-class Recall	0.601312404421
Binary F-score	0.908264963828
Binary Precision	0.840760804501
Binary Recall	0.987555149321
Total amount of samples	195519
False negative	708
False positive	10641
True negative	127987
True positive	56183
Positive training samples	25195
Negative training samples	24806

Table 4.2: DecisionTreeClassifier with Standard feature set CTU:
Experiment 2.

Multi-class F-score	0.599723082938
Multi-class Precision	0.656979098165
Multi-class Recall	0.60028437134
Binary F-score	0.906333688252
Binary Precision	0.837494968769
Binary Recall	0.987502416902
Total amount of samples	195519
False negative	711
False positive	10901
True negative	127727
True positive	56180
Positive training samples	25195
Negative training samples	24806

Table 4.3: DecisionTreeClassifier with Standard feature set CTU:
Experiment 3.

Multi-class F-score	0.599966992355
Multi-class Precision	0.656486962224
Multi-class Recall	0.600949268358
Binary F-score	0.907934712346
Binary Precision	0.84019500815
Binary Recall	0.987555149321
Total amount of samples	195519
False negative	708
False positive	10686
True negative	127942
True positive	56183
Positive training samples	25195
Negative training samples	24806

Table 4.4: DecisionTreeClassifier with Standard feature set CTU:
Experiment 4.

Multi-class F-score	0.599850550114
Multi-class Precision	0.656920347771
Multi-class Recall	0.600545215554
Binary F-score	0.906686248245
Binary Precision	0.838084582681
Binary Recall	0.987519994375
Total amount of samples	195519
False negative	710
False positive	10854
True negative	127774
True positive	56181
Positive training samples	25195
Negative training samples	24806

Table 4.5: DecisionTreeClassifier with Standard feature set CTU:
Experiment 5.

Multi-class F-score	0.599608827077
Multi-class Precision	0.657068251333
Multi-class Recall	0.600330402672
Binary F-score	0.90757388516
Binary Precision	0.839589933648
Binary Recall	0.987537571848
Total amount of samples	195519
False negative	709
False positive	10734
True negative	127894
True positive	56182
Positive training samples	25195
Negative training samples	24806

Table 4.6: DecisionTreeClassifier with Standard feature set CTU:
Experiment 6.

Multi-class F-score	0.600791622873
Multi-class Precision	0.657826758729
Multi-class Recall	0.60162439456
Binary F-score	0.907474258025
Binary Precision	0.83939402994
Binary Recall	0.987572726793
Total amount of samples	195519
False negative	707
False positive	10750
True negative	127878
True positive	56184
Positive training samples	25195
Negative training samples	24806

Table 4.7: DecisionTreeClassifier with Standard feature set CTU:
Experiment 7.

Multi-class F-score	0.599833298642
Multi-class Precision	0.656600037906
Multi-class Recall	0.60070376792
Binary F-score	0.907686714869
Binary Precision	0.839821209675
Binary Recall	0.98748483943
Total amount of samples	195519
False negative	712
False positive	10715
True negative	127913
True positive	56179
Positive training samples	25195
Negative training samples	24806

Table 4.8: DecisionTreeClassifier with Standard feature set CTU:
Experiment 8.

Multi-class F-score	0.600420916922
Multi-class Precision	0.65744375654
Multi-class Recall	0.601092476946
Binary F-score	0.907184033328
Binary Precision	0.838935595144
Binary Recall	0.987519994375
Total amount of samples	195519
False negative	710
False positive	10786
True negative	127842
True positive	56181
Positive training samples	25195
Negative training samples	24806

Table 4.9: DecisionTreeClassifier with Standard feature set CTU:
Experiment 9.

Multi-class F-score	0.599721499396
Multi-class Precision	0.656800586978
Multi-class Recall	0.600412236151
Binary F-score	0.907408304666
Binary Precision	0.839281179511
Binary Recall	0.987572726793
Total amount of samples	195519
False negative	707
False positive	10759
True negative	127869
True positive	56184
Positive training samples	25195
Negative training samples	24806

Table 4.10: DecisionTreeClassifier with Standard feature set CTU:
Experiment 10.

Multi-class F-score	0.600119012652
Multi-class Precision	0.657045853748
Multi-class Recall	0.600826518139
Binary F-score	0.907100649036
Binary Precision	0.838754945137
Binary Recall	0.987572726793
Total amount of samples	195519
False negative	707
False positive	10801
True negative	127827
True positive	56184
Positive training samples	25195
Negative training samples	24806

Table 4.11: DecisionTreeClassifier with
Standard feature set CTU:
Average.

Multi-class F-score	0.600058640709
Multi-class Precision	0.657093970884
Multi-class Recall	0.600808105606
Binary F-score	0.907364745776
Binary Precision	0.839231225716
Binary Recall	0.987539329595
Total amount of samples	195519.0
False negative	708.9
False positive	10762.7
True negative	127865.3
True positive	56182.1
Positive training samples	25195.0
Negative training samples	24806.0

Table 4.12: DecisionTreeClassifier with
Standard feature set CTU:
Variance.

Multi-class F-score	1.44525542814e-07
Multi-class Precision	1.85853892895e-07
Multi-class Recall	1.76098578635e-07
Binary F-score	2.92239071481e-07
Binary Precision	8.41785995193e-07
Binary Recall	8.92916210555e-10
Total amount of samples	0.0
False negative	2.89
False positive	5344.01
True negative	5344.01
True positive	2.89
Positive training samples	0.0
Negative training samples	0.0

4.1.2 Cross Dataset

Table 4.13: DecisionTreeClassifier with
Standard feature set Cross:
Experiment 1.

Multi-class F-score	0.751456683266
Multi-class Precision	0.789500071853
Multi-class Recall	0.745856443256
Binary F-score	0.94128452462
Binary Precision	0.894592201544
Binary Recall	0.993119372398
Total amount of samples	246467
False negative	742
False positive	12619
True negative	126009
True positive	107097
Positive training samples	30140
Negative training samples	19861

Table 4.14: DecisionTreeClassifier with
Standard feature set Cross:
Experiment 2.

Multi-class F-score	0.751175623771
Multi-class Precision	0.789881997219
Multi-class Recall	0.745568372236
Binary F-score	0.941342930085
Binary Precision	0.894569810753
Binary Recall	0.993277014809
Total amount of samples	246467
False negative	725
False positive	12624
True negative	126004
True positive	107114
Positive training samples	30140
Negative training samples	19861

Table 4.15: DecisionTreeClassifier with
Standard feature set Cross:
Experiment 3.

Multi-class F-score	0.750989198786
Multi-class Precision	0.789976506307
Multi-class Recall	0.745191039774
Binary F-score	0.941139774365
Binary Precision	0.894210460408
Binary Recall	0.993267741726
Total amount of samples	246467
False negative	726
False positive	12672
True negative	125956
True positive	107113
Positive training samples	30140
Negative training samples	19861

Table 4.16: DecisionTreeClassifier with
Standard feature set Cross:
Experiment 4.

Multi-class F-score	0.750398146373
Multi-class Precision	0.789322974362
Multi-class Recall	0.74457838169
Binary F-score	0.941081886118
Binary Precision	0.894166068358
Binary Recall	0.993193557062
Total amount of samples	246467
False negative	734
False positive	12677
True negative	125951
True positive	107105
Positive training samples	30140
Negative training samples	19861

Table 4.17: DecisionTreeClassifier with Standard feature set Cross: Experiment 5.

Multi-class F-score	0.750557597117
Multi-class Precision	0.788966819949
Multi-class Recall	0.74505714761
Binary F-score	0.941684247539
Binary Precision	0.895096130547
Binary Recall	0.993388291805
Total amount of samples	246467
False negative	713
False positive	12555
True negative	126073
True positive	107126
Positive training samples	30140
Negative training samples	19861

Table 4.18: DecisionTreeClassifier with Standard feature set Cross: Experiment 6.

Multi-class F-score	0.750731991378
Multi-class Precision	0.789012983495
Multi-class Recall	0.745199154451
Binary F-score	0.941428791175
Binary Precision	0.894800160406
Binary Recall	0.993184283979
Total amount of samples	246467
False negative	735
False positive	12592
True negative	126036
True positive	107104
Positive training samples	30140
Negative training samples	19861

Table 4.19: DecisionTreeClassifier with Standard feature set Cross: Experiment 7.

Multi-class F-score	0.751210922617
Multi-class Precision	0.790376519046
Multi-class Recall	0.745576486913
Binary F-score	0.941279348242
Binary Precision	0.894537709847
Binary Recall	0.993175010896
Total amount of samples	246467
False negative	736
False positive	12627
True negative	126001
True positive	107103
Positive training samples	30140
Negative training samples	19861

Table 4.20: DecisionTreeClassifier with Standard feature set Cross: Experiment 8.

Multi-class F-score	0.751183766726
Multi-class Precision	0.789438228138
Multi-class Recall	0.745625174973
Binary F-score	0.94168739945
Binary Precision	0.89517712834
Binary Recall	0.993295560975
Total amount of samples	246467
False negative	723
False positive	12543
True negative	126085
True positive	107116
Positive training samples	30140
Negative training samples	19861

Table 4.21: DecisionTreeClassifier with Standard feature set Cross: Experiment 9.

Multi-class F-score	0.750473561247
Multi-class Precision	0.789039920758
Multi-class Recall	0.744878624725
Binary F-score	0.941773910597
Binary Precision	0.895461614354
Binary Recall	0.993137918564
Total amount of samples	246467
False negative	740
False positive	12503
True negative	126125
True positive	107099
Positive training samples	30140
Negative training samples	19861

Table 4.22: DecisionTreeClassifier with Standard feature set Cross: Experiment 10.

Multi-class F-score	0.750513737741
Multi-class Precision	0.789215825137
Multi-class Recall	0.744886739401
Binary F-score	0.940843734355
Binary Precision	0.893653587846
Binary Recall	0.993295560975
Total amount of samples	246467
False negative	723
False positive	12747
True negative	125881
True positive	107116
Positive training samples	30140
Negative training samples	19861

Table 4.23: DecisionTreeClassifier with Standard feature set Cross: Average.

Multi-class F-score	0.750869122902
Multi-class Precision	0.789473184626
Multi-class Recall	0.745241756503
Binary F-score	0.941354654654
Binary Precision	0.89462648724
Binary Recall	0.993233431319
Total amount of samples	246467.0
False negative	729.7
False positive	12615.9
True negative	126012.1
True positive	107109.3
Positive training samples	30140.0
Negative training samples	19861.0

Table 4.24: DecisionTreeClassifier with Standard feature set Cross: Variance.

Multi-class F-score	1.28985585938e-07
Multi-class Precision	1.98312328034e-07
Multi-class Recall	1.48000735042e-07
Binary F-score	7.92964846173e-08
Binary Precision	2.59523638352e-07
Binary Recall	6.43291704179e-09
Total amount of samples	0.0
False negative	74.81
False positive	4660.69
True negative	4660.69
True positive	74.81
Positive training samples	0.0
Negative training samples	0.0

4.2 TCP feature set

4.2.1 CTU Dataset

Table 4.25: DecisionTreeClassifier with TCP feature set CTU: Experiment 1.

Multi-class F-score	0.664852721831
Multi-class Precision	0.709460994347
Multi-class Recall	0.673044563444
Binary F-score	0.950852421826
Binary Precision	0.917068073086
Binary Recall	0.987221177339
Total amount of samples	195519
False negative	18600.8217411
False positive	15680.6229042
True negative	122947.377096
True positive	38290.1782589
Positive training samples	25195
Negative training samples	24806

Table 4.26: DecisionTreeClassifier with TCP feature set CTU: Experiment 2.

Multi-class F-score	0.66502260591
Multi-class Precision	0.709551737503
Multi-class Recall	0.673474189209
Binary F-score	0.95141998561
Binary Precision	0.917608255503
Binary Recall	0.987818811411
Total amount of samples	195519
False negative	18576.3799017
False positive	15683.7257209
True negative	122944.274279
True positive	38314.6200983
Positive training samples	25195
Negative training samples	24806

Table 4.27: DecisionTreeClassifier with TCP feature set CTU: Experiment 3.

Multi-class F-score	0.665331534955
Multi-class Precision	0.709902731669
Multi-class Recall	0.673596939428
Binary F-score	0.951201747997
Binary Precision	0.917809338585
Binary Recall	0.987115712503
Total amount of samples	195519
False negative	18569.396519
False positive	15659.8812654
True negative	122968.118735
True positive	38321.603481
Positive training samples	25195
Negative training samples	24806

Table 4.28: DecisionTreeClassifier with TCP feature set CTU: Experiment 4.

Multi-class F-score	0.665707088832
Multi-class Precision	0.710513728026
Multi-class Recall	0.674057252748
Binary F-score	0.951557737894
Binary Precision	0.917849385095
Binary Recall	0.987836388884
Total amount of samples	195519
False negative	18543.2088339
False positive	15624.1303514
True negative	123003.869649
True positive	38347.7911661
Positive training samples	25195
Negative training samples	24806

Table 4.29: DecisionTreeClassifier with TCP feature set CTU: Experiment 5.

Multi-class F-score	0.664508568392
Multi-class Precision	0.709257808118
Multi-class Recall	0.672865552708
Binary F-score	0.951093210587
Binary Precision	0.916939944202
Binary Recall	0.987889121302
Total amount of samples	195519
False negative	18611.0058409
False positive	15691.9095985
True negative	122936.090401
True positive	38279.9941591
Positive training samples	25195
Negative training samples	24806

Table 4.30: DecisionTreeClassifier with TCP feature set CTU: Experiment 6.

Multi-class F-score	0.664907883302
Multi-class Precision	0.710152429256
Multi-class Recall	0.673090594776
Binary F-score	0.950846596681
Binary Precision	0.917163435627
Binary Recall	0.98709813503
Total amount of samples	195519
False negative	18598.2029726
False positive	15629.1434601
True negative	122998.85654
True positive	38292.7970274
Positive training samples	25195
Negative training samples	24806

Table 4.31: DecisionTreeClassifier with TCP feature set CTU: Experiment 7.

Multi-class F-score	0.665000086504
Multi-class Precision	0.709951546627
Multi-class Recall	0.673346324398
Binary F-score	0.951167872085
Binary Precision	0.917776651903
Binary Recall	0.987080557557
Total amount of samples	195519
False negative	18583.6542587
False positive	15650.3446438
True negative	122977.655356
True positive	38307.3457413
Positive training samples	25195
Negative training samples	24806

Table 4.32: DecisionTreeClassifier with TCP feature set CTU: Experiment 8.

Multi-class F-score	0.664918454515
Multi-class Precision	0.709943913238
Multi-class Recall	0.673284949289
Binary F-score	0.951146408372
Binary Precision	0.917539245634
Binary Recall	0.987309064703
Total amount of samples	195519
False negative	18587.14595
False positive	15649.49823
True negative	122978.50177
True positive	38303.85405
Positive training samples	25195
Negative training samples	24806

Table 4.33: DecisionTreeClassifier with TCP feature set CTU: Experiment 9.

Multi-class F-score	0.664639419081
Multi-class Precision	0.710319278269
Multi-class Recall	0.672845094339
Binary F-score	0.95107618838
Binary Precision	0.917514825767
Binary Recall	0.987186022394
Total amount of samples	195519
False negative	18612.169738
False positive	15610.7816817
True negative	123017.218318
True positive	38278.830262
Positive training samples	25195
Negative training samples	24806

Table 4.34: DecisionTreeClassifier with TCP feature set CTU: Experiment 10.

Multi-class F-score	0.665780933583
Multi-class Precision	0.71069155127
Multi-class Recall	0.67410328408
Binary F-score	0.950943779692
Binary Precision	0.91748096343
Binary Recall	0.986939937776
Total amount of samples	195519
False negative	18540.5900654
False positive	15611.6920013
True negative	123016.307999
True positive	38350.4099346
Positive training samples	25195
Negative training samples	24806

Table 4.35: DecisionTreeClassifier with TCP feature set CTU: Average.

Multi-class F-score	0.665066929691
Multi-class Precision	0.709974571832
Multi-class Recall	0.673370874442
Binary F-score	0.951130594912
Binary Precision	0.917475011883
Binary Recall	0.98734949289
Total amount of samples	195519
False negative	18582.2575821
False positive	15649.1729857
True negative	122978.827014
True positive	38308.7424179
Positive training samples	25195.0
Negative training samples	24806.0

Table 4.36: DecisionTreeClassifier with TCP feature set CTU: Variance.

Multi-class F-score	1.58381213855e-07
Multi-class Precision	1.91813674992e-07
Multi-class Recall	1.79413454195e-07
Binary F-score	4.70193363196e-08
Binary Precision	9.1539312147e-08
Binary Recall	1.15124397652e-07
Total amount of samples	195519
False negative	580.687052883
False positive	803.258463249
True negative	803.258463249
True positive	580.687052883
Positive training samples	0.0
Negative training samples	0.0

4.2.2 Cross Dataset

Table 4.37: DecisionTreeClassifier with TCP feature set Cross: Experiment 1.

Multi-class F-score	0.804074298918
Multi-class Precision	0.830762578285
Multi-class Recall	0.805596692458
Binary F-score	0.971247483815
Binary Precision	0.950165435143
Binary Recall	0.993286287892
Total amount of samples	246467
False negative	20964.258282
False positive	17697.5440214
True negative	120930.455979
True positive	86874.741718
Positive training samples	30140
Negative training samples	19861

Table 4.38: DecisionTreeClassifier with TCP feature set Cross: Experiment 2.

Multi-class F-score	0.803638357068
Multi-class Precision	0.830210640663
Multi-class Recall	0.805085467831
Binary F-score	0.971677915669
Binary Precision	0.951032186459
Binary Recall	0.993239922477
Total amount of samples	246467
False negative	21019.3882346
False positive	17755.7905639
True negative	120872.209436
True positive	86819.6117654
Positive training samples	30140
Negative training samples	19861

Table 4.39: DecisionTreeClassifier with TCP feature set Cross: Experiment 3.

Multi-class F-score	0.804955096221
Multi-class Precision	0.831186777151
Multi-class Recall	0.806935614098
Binary F-score	0.972681162376
Binary Precision	0.951382361808
Binary Recall	0.994955442836
Total amount of samples	246467
False negative	20819.8703113
False positive	17673.5002722
True negative	120954.499728
True positive	87019.1296887
Positive training samples	30140
Negative training samples	19861

Table 4.40: DecisionTreeClassifier with TCP feature set Cross: Experiment 4.

Multi-class F-score	0.804199864239
Multi-class Precision	0.830049639495
Multi-class Recall	0.805685953901
Binary F-score	0.971548826894
Binary Precision	0.950827385391
Binary Recall	0.993193557062
Total amount of samples	246467
False negative	20954.6324173
False positive	17789.33318
True negative	120838.66682
True positive	86884.3675827
Positive training samples	30140
Negative training samples	19861

Table 4.41: DecisionTreeClassifier with TCP feature set Cross: Experiment 5.

Multi-class F-score	0.804361653737
Multi-class Precision	0.830825653429
Multi-class Recall	0.805994311612
Binary F-score	0.971149051147
Binary Precision	0.95002793965
Binary Recall	0.993230649394
Total amount of samples	246467
False negative	20921.3794301
False positive	17698.3361127
True negative	120929.663887
True positive	86917.6205699
Positive training samples	30140
Negative training samples	19861

Table 4.42: DecisionTreeClassifier with TCP feature set Cross: Experiment 6.

Multi-class F-score	0.804860765836
Multi-class Precision	0.831329370479
Multi-class Recall	0.806375701412
Binary F-score	0.970919902825
Binary Precision	0.949597936042
Binary Recall	0.993221376311
Total amount of samples	246467
False negative	20916.1289585
False positive	17714.7224561
True negative	120913.277544
True positive	86922.8710415
Positive training samples	30140
Negative training samples	19861

Table 4.43: DecisionTreeClassifier with TCP feature set Cross: Experiment 7.

Multi-class F-score	0.804663432405
Multi-class Precision	0.830704034143
Multi-class Recall	0.806042999671
Binary F-score	0.971828941998
Binary Precision	0.950990050679
Binary Recall	0.993601572715
Total amount of samples	246467
False negative	20880.2507354
False positive	17643.2921796
True negative	120984.70782
True positive	86958.7492646
Positive training samples	30140
Negative training samples	19861

Table 4.44: DecisionTreeClassifier with TCP feature set Cross: Experiment 8.

Multi-class F-score	0.804502267242
Multi-class Precision	0.830998233308
Multi-class Recall	0.805815788726
Binary F-score	0.97163557885
Binary Precision	0.950602826498
Binary Recall	0.993620118881
Total amount of samples	246467
False negative	20940.6311596
False positive	17672.6944391
True negative	120955.305561
True positive	86898.3688404
Positive training samples	30140
Negative training samples	19861

Table 4.45: DecisionTreeClassifier with TCP feature set Cross: Experiment 9.

Multi-class F-score	0.803759498349
Multi-class Precision	0.831152345488
Multi-class Recall	0.804955633006
Binary F-score	0.971655292811
Binary Precision	0.950716973096
Binary Recall	0.993536661134
Total amount of samples	246467
False negative	21033.3894923
False positive	17634.4611337
True negative	120993.538866
True positive	86805.6105077
Positive training samples	30140
Negative training samples	19861

Table 4.46: DecisionTreeClassifier with TCP feature set Cross: Experiment 10.

Multi-class F-score	0.803826097192
Multi-class Precision	0.830129376312
Multi-class Recall	0.805332965468
Binary F-score	0.971786919142
Binary Precision	0.950841637311
Binary Recall	0.993675757379
Total amount of samples	246467
False negative	20992.6983369
False positive	17771.4894202
True negative	120856.51058
True positive	86846.3016631
Positive training samples	30140
Negative training samples	19861

Table 4.47: DecisionTreeClassifier with TCP feature set Cross: Average.

Multi-class F-score	0.804284133121
Multi-class Precision	0.830734864875
Multi-class Recall	0.805782112818
Binary F-score	0.971613107553
Binary Precision	0.950618473208
Binary Recall	0.993556134608
Total amount of samples	246467
False negative	20944.2627358
False positive	17705.1163779
True negative	120922.883622
True positive	86894.7372642
Positive training samples	30140.0
Negative training samples	19861.0

Table 4.48: DecisionTreeClassifier with TCP feature set Cross: Variance.

Multi-class F-score	1.93335356128e-07
Multi-class Precision	1.92223698782e-07
Multi-class Recall	3.21088067544e-07
Binary F-score	2.05771459414e-07
Binary Precision	2.59161730669e-07
Binary Recall	2.50118453018e-07
Total amount of samples	246467
False negative	3734.0133821
False positive	2512.64682287
True negative	2512.64682287
True positive	3734.0133821
Positive training samples	0.0
Negative training samples	0.0

4.3 Country feature set

4.3.1 CTU Dataset

Table 4.49: DecisionTreeClassifier with Country feature set CTU: Experiment 1.

Multi-class F-score	0.599977008976
Multi-class Precision	0.65682599607
Multi-class Recall	0.600662851181
Binary F-score	0.906866187236
Binary Precision	0.838417455153
Binary Recall	0.98748483943
Total amount of samples	195519
False negative	712
False positive	10827
True negative	127801
True positive	56179
Positive training samples	25195
Negative training samples	24806

Table 4.50: DecisionTreeClassifier with Country feature set CTU: Experiment 2.

Multi-class F-score	0.60004990241
Multi-class Precision	0.656948151542
Multi-class Recall	0.600770257622
Binary F-score	0.90740082698
Binary Precision	0.839331869183
Binary Recall	0.98748483943
Total amount of samples	195519
False negative	712
False positive	10754
True negative	127874
True positive	56179
Positive training samples	25195
Negative training samples	24806

Table 4.51: DecisionTreeClassifier with Country feature set CTU: Experiment 3.

Multi-class F-score	0.599917375249
Multi-class Precision	0.65699856449
Multi-class Recall	0.60072422629
Binary F-score	0.907386322343
Binary Precision	0.839243569444
Binary Recall	0.987572726793
Total amount of samples	195519
False negative	707
False positive	10762
True negative	127866
True positive	56184
Positive training samples	25195
Negative training samples	24806

Table 4.52: DecisionTreeClassifier with Country feature set CTU: Experiment 4.

Multi-class F-score	0.598766880772
Multi-class Precision	0.656271404448
Multi-class Recall	0.599527411658
Binary F-score	0.906914614559
Binary Precision	0.838462227263
Binary Recall	0.987537571848
Total amount of samples	195519
False negative	709
False positive	10824
True negative	127804
True positive	56182
Positive training samples	25195
Negative training samples	24806

Table 4.53: DecisionTreeClassifier with Country feature set CTU: Experiment 5.

Multi-class F-score	0.599891242676
Multi-class Precision	0.657164575279
Multi-class Recall	0.600632163626
Binary F-score	0.906573501053
Binary Precision	0.837917250843
Binary Recall	0.98748483943
Total amount of samples	195519
False negative	712
False positive	10867
True negative	127761
True positive	56179
Positive training samples	25195
Negative training samples	24806

Table 4.54: DecisionTreeClassifier with Country feature set CTU: Experiment 6.

Multi-class F-score	0.599938281319
Multi-class Precision	0.657061863905
Multi-class Recall	0.600760028437
Binary F-score	0.906740475941
Binary Precision	0.838151927438
Binary Recall	0.987555149321
Total amount of samples	195519
False negative	708
False positive	10849
True negative	127779
True positive	56183
Positive training samples	25195
Negative training samples	24806

Table 4.55: DecisionTreeClassifier with Country feature set CTU: Experiment 7.

Multi-class F-score	0.600865186318
Multi-class Precision	0.657928901797
Multi-class Recall	0.601629509153
Binary F-score	0.907192856393
Binary Precision	0.838938000239
Binary Recall	0.987537571848
Total amount of samples	195519
False negative	709
False positive	10786
True negative	127842
True positive	56182
Positive training samples	25195
Negative training samples	24806

Table 4.56: DecisionTreeClassifier with Country feature set CTU: Experiment 8.

Multi-class F-score	0.599746659081
Multi-class Precision	0.656946729948
Multi-class Recall	0.60047361126
Binary F-score	0.906212447477
Binary Precision	0.837262674734
Binary Recall	0.987537571848
Total amount of samples	195519
False negative	709
False positive	10920
True negative	127708
True positive	56182
Positive training samples	25195
Negative training samples	24806

Table 4.57: DecisionTreeClassifier with
Country feature set CTU:
Experiment 9.

Multi-class F-score	0.600335687345
Multi-class Precision	0.657389604878
Multi-class Recall	0.601240800127
Binary F-score	0.909014417242
Binary Precision	0.842109996852
Binary Recall	0.987467261957
Total amount of samples	195519
False negative	713
False positive	10533
True negative	128095
True positive	56178
Positive training samples	25195
Negative training samples	24806

Table 4.58: DecisionTreeClassifier with
Country feature set CTU:
Experiment 10.

Multi-class F-score	0.599900321173
Multi-class Precision	0.656653704741
Multi-class Recall	0.600729340882
Binary F-score	0.908021174284
Binary Precision	0.840419484172
Binary Recall	0.987449684484
Total amount of samples	195519
False negative	714
False positive	10667
True negative	127961
True positive	56177
Positive training samples	25195
Negative training samples	24806

Table 4.59: DecisionTreeClassifier with
Country feature set CTU:
Average.

Multi-class F-score	0.599938854532
Multi-class Precision	0.657019078905
Multi-class Recall	0.600715020024
Binary F-score	0.907232282351
Binary Precision	0.839025445532
Binary Recall	0.987511205639
Total amount of samples	195519.0
False negative	710.5
False positive	10778.9
True negative	127849.1
True positive	56180.5
Positive training samples	25195.0
Negative training samples	24806.0

Table 4.60: DecisionTreeClassifier with
Country feature set CTU:
Variance.

Multi-class F-score	2.44402923375e-07
Multi-class Precision	1.72849786148e-07
Multi-class Recall	2.59638046722e-07
Binary F-score	5.80291324381e-07
Binary Precision	1.73922780861e-06
Binary Recall	1.56028611187e-09
Total amount of samples	0.0
False negative	5.05
False positive	11087.69
True negative	11087.69
True positive	5.05
Positive training samples	0.0
Negative training samples	0.0

4.3.2 Cross Dataset

Table 4.61: DecisionTreeClassifier with Country feature set Cross: Experiment 1.

Multi-class F-score	0.750514171292
Multi-class Precision	0.78895469205
Multi-class Recall	0.744882682063
Binary F-score	0.941782056354
Binary Precision	0.895295484367
Binary Recall	0.993360472556
Total amount of samples	246467
False negative	716
False positive	12528
True negative	126100
True positive	107123
Positive training samples	30140
Negative training samples	19861

Table 4.62: DecisionTreeClassifier with Country feature set Cross: Experiment 2.

Multi-class F-score	0.750462077799
Multi-class Precision	0.789343029936
Multi-class Recall	0.744785305944
Binary F-score	0.941134608457
Binary Precision	0.894156044708
Binary Recall	0.993323380224
Total amount of samples	246467
False negative	720
False positive	12680
True negative	125948
True positive	107119
Positive training samples	30140
Negative training samples	19861

Table 4.63: DecisionTreeClassifier with Country feature set Cross: Experiment 3.

Multi-class F-score	0.750736467976
Multi-class Precision	0.789479693438
Multi-class Recall	0.745170753083
Binary F-score	0.941742750263
Binary Precision	0.895254571591
Binary Recall	0.993323380224
Total amount of samples	246467
False negative	720
False positive	12533
True negative	126095
True positive	107119
Positive training samples	30140
Negative training samples	19861

Table 4.64: DecisionTreeClassifier with Country feature set Cross: Experiment 4.

Multi-class F-score	0.751086165168
Multi-class Precision	0.789751164854
Multi-class Recall	0.745479110794
Binary F-score	0.941109798047
Binary Precision	0.894171369419
Binary Recall	0.99324919556
Total amount of samples	246467
False negative	728
False positive	12677
True negative	125951
True positive	107111
Positive training samples	30140
Negative training samples	19861

Table 4.65: DecisionTreeClassifier with Country feature set Cross: Experiment 5.

Multi-class F-score	0.750781804033
Multi-class Precision	0.789592753592
Multi-class Recall	0.744976000844
Binary F-score	0.941400311102
Binary Precision	0.894620800241
Binary Recall	0.99334192639
Total amount of samples	246467
False negative	718
False positive	12618
True negative	126010
True positive	107121
Positive training samples	30140
Negative training samples	19861

Table 4.66: DecisionTreeClassifier with Country feature set Cross: Experiment 6.

Multi-class F-score	0.750916646913
Multi-class Precision	0.789199120926
Multi-class Recall	0.745365505321
Binary F-score	0.941574683389
Binary Precision	0.89510898636
Binary Recall	0.993128645481
Total amount of samples	246467
False negative	741
False positive	12550
True negative	126078
True positive	107098
Positive training samples	30140
Negative training samples	19861

Table 4.67: DecisionTreeClassifier with Country feature set Cross: Experiment 7.

Multi-class F-score	0.750618223404
Multi-class Precision	0.789042220241
Multi-class Recall	0.744980058182
Binary F-score	0.941577735192
Binary Precision	0.89500906607
Binary Recall	0.993258468643
Total amount of samples	246467
False negative	727
False positive	12565
True negative	126063
True positive	107112
Positive training samples	30140
Negative training samples	19861

Table 4.68: DecisionTreeClassifier with Country feature set Cross: Experiment 8.

Multi-class F-score	0.751045124812
Multi-class Precision	0.78958356398
Multi-class Recall	0.745292473232
Binary F-score	0.941209032598
Binary Precision	0.894290413338
Binary Recall	0.993323380224
Total amount of samples	246467
False negative	720
False positive	12662
True negative	125966
True positive	107119
Positive training samples	30140
Negative training samples	19861

Table 4.69: DecisionTreeClassifier with Country feature set Cross: Experiment 9.

Multi-class F-score	0.751196086467
Multi-class Precision	0.789576620854
Multi-class Recall	0.745779353828
Binary F-score	0.941414816149
Binary Precision	0.894737281764
Binary Recall	0.993230649394
Total amount of samples	246467
False negative	730
False positive	12601
True negative	126027
True positive	107109
Positive training samples	30140
Negative training samples	19861

Table 4.70: DecisionTreeClassifier with Country feature set Cross: Experiment 10.

Multi-class F-score	0.750223087693
Multi-class Precision	0.788658249746
Multi-class Recall	0.74457838169
Binary F-score	0.940738398848
Binary Precision	0.893433530447
Binary Recall	0.993332653307
Total amount of samples	246467
False negative	719
False positive	12777
True negative	125851
True positive	107120
Positive training samples	30140
Negative training samples	19861

Table 4.71: DecisionTreeClassifier with Country feature set Cross: Average.

Multi-class F-score	0.750757985556
Multi-class Precision	0.789318110962
Multi-class Recall	0.745128962498
Binary F-score	0.94136841904
Binary Precision	0.89460775483
Binary Recall	0.9932872152
Total amount of samples	246467.0
False negative	723.9
False positive	12619.1
True negative	126008.9
True positive	107115.1
Positive training samples	30140.0
Negative training samples	19861.0

Table 4.72: DecisionTreeClassifier with Country feature set Cross: Variance.

Multi-class F-score	8.60942949731e-08
Multi-class Precision	1.08475121707e-07
Multi-class Recall	1.1575103008e-07
Binary F-score	9.44629407473e-08
Binary Precision	3.19455779064e-07
Binary Recall	4.49642069396e-09
Total amount of samples	0.0
False negative	52.29
False positive	5757.69
True negative	5757.69
True positive	52.29
Positive training samples	0.0
Negative training samples	0.0

Chapter 5

Naive Bayes

Chapter 6

Support Vector Machines with Linear Kernel

Chapter 7

Support Vector Machine with RBF Kernel

Chapter 8

Neural Networks

Chapter 9

One class Support Vector Machines