Network Intrusion Detection

Learning & Softcomputing

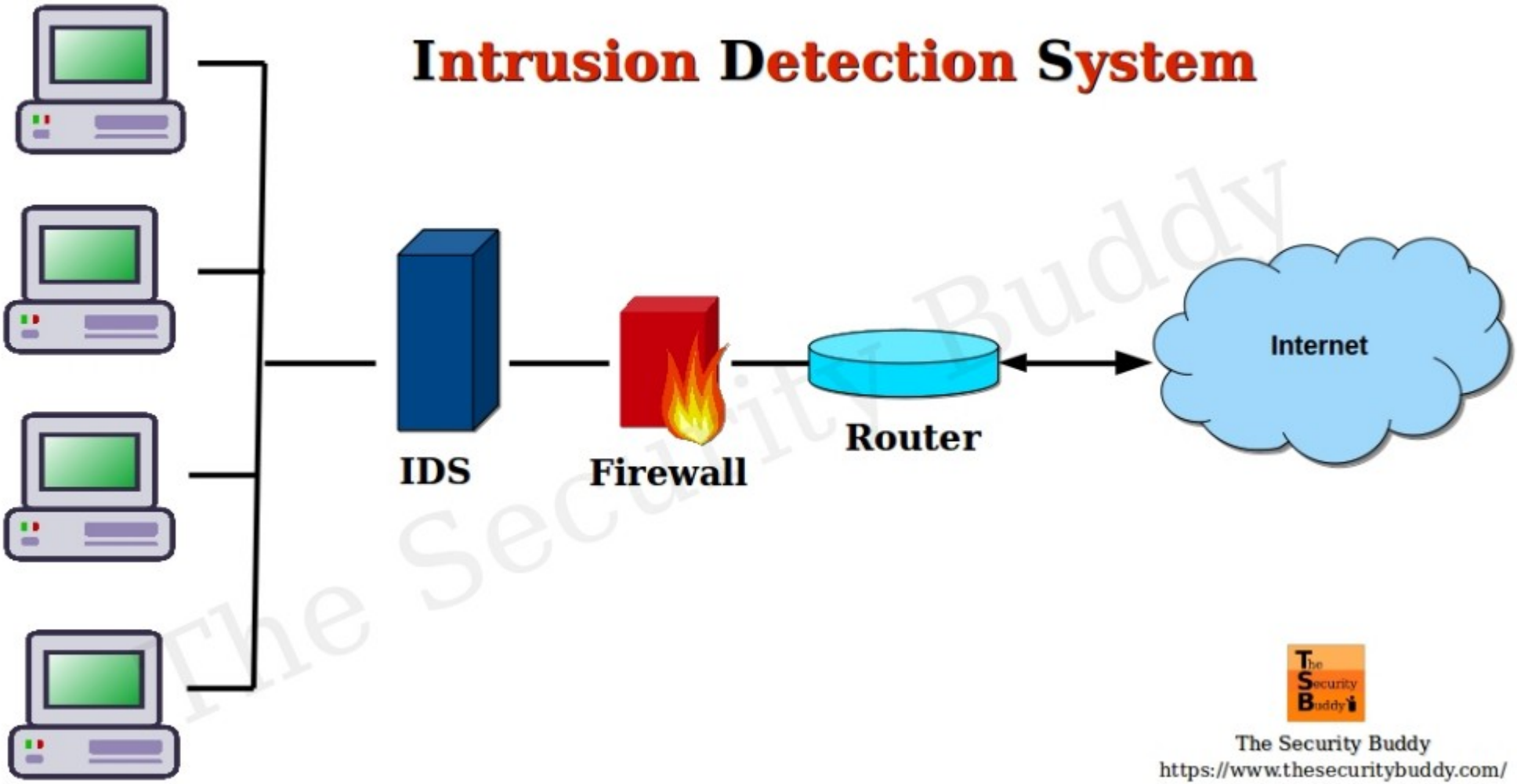
Alexander Belzer
Matr. Nr.: inf104862

Maximilian Wendt
Matr. Nr. its105814

Advisor:
Prof. Dr. Hoffmann

Roadmap

- 1. Einführung
2. Datensatz
3. Verfolgter Ansatz
4. Ergebnisse



Source: https://www.thesecuritybuddy.com/data-breaches-prevention/what-is-ids-intrusion-detection-system-how-does-it-work/

Roadmap

1. Einführung
2. Datensatz
    - 2.1 NSL KDD
    - 2.2 Besonderheiten
    - 2.3 Bereinigung des Datensatzes
    - 2.4 Alternativen
3. Verfolgter Ansatz
4. Ergebnisse

Weiterentwicklung des KDD 1999 Datensatzes

Aufteilung in Trainings- und Testdaten bereits erfolgt.

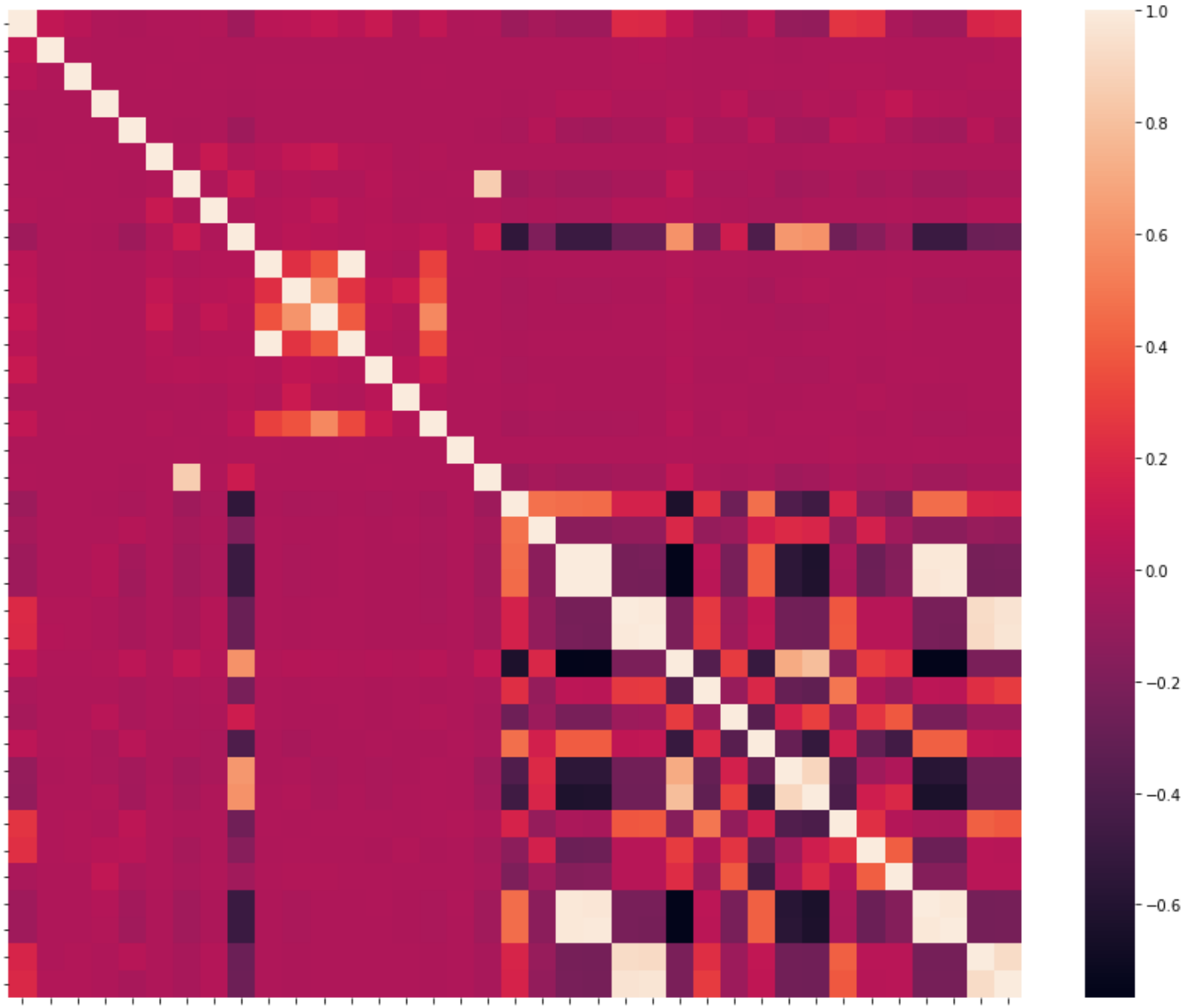
Vier Kategorien von Attacken:

DOS
Probe
R2L
U2L

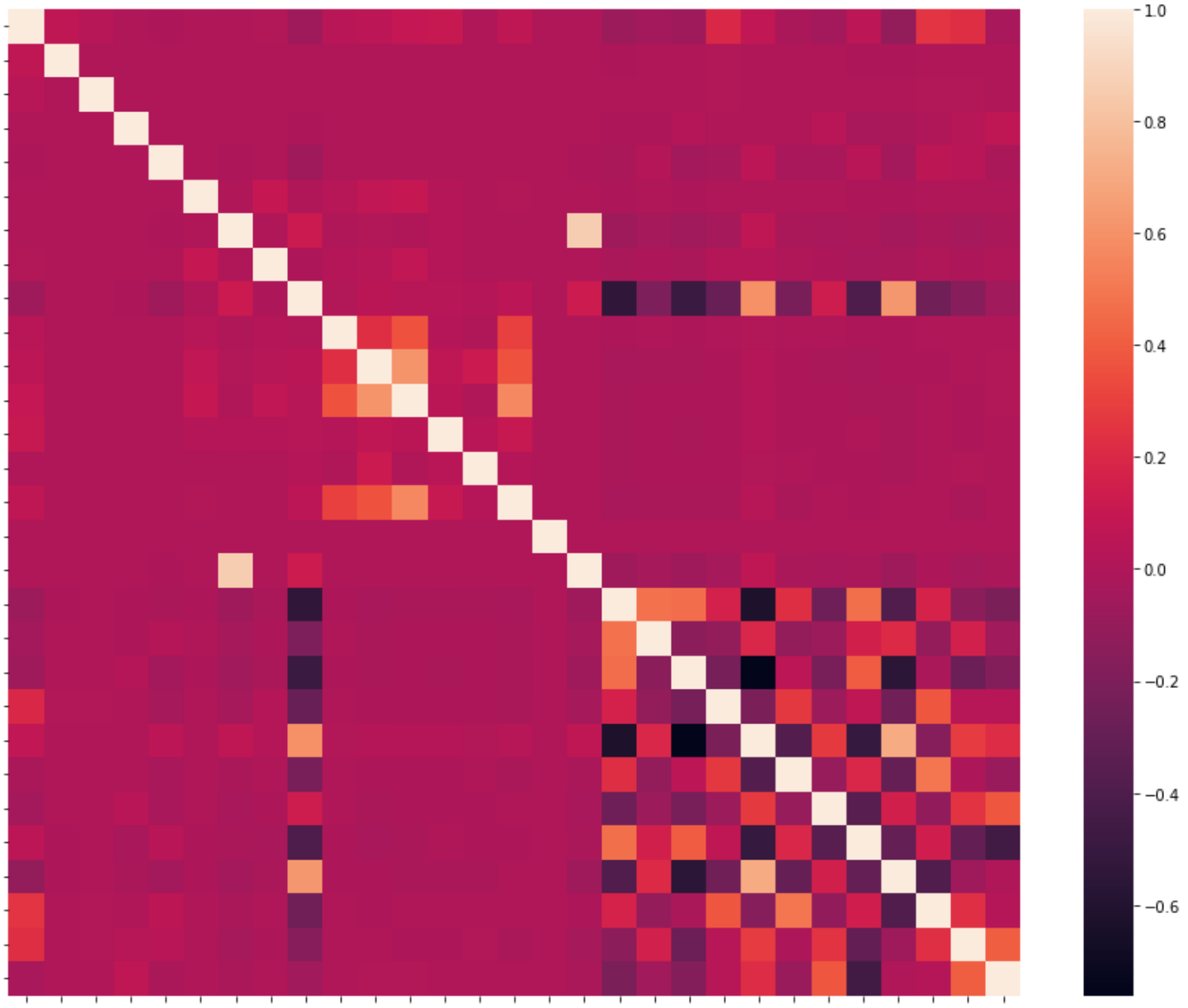
2.3 Bereinigung des Datensatzes

	count	mean	std	min	25%	50%	75%	max
duration	125973.0	287.144650	2.604515e+03	0.0	0.00	0.00	0.00	4.290800e+04
src_bytes	125973.0	45566.743000	5.870331e+06	0.0	0.00	44.00	276.00	1.379964e+09
dst_bytes	125973.0	19779.114421	4.021269e+06	0.0	0.00	0.00	516.00	1.309937e+09
land	125973.0	0.000198	1.408607e-02	0.0	0.00	0.00	0.00	1.000000e+00
wrong_fragment	125973.0	0.022687	2.535300e-01	0.0	0.00	0.00	0.00	3.000000e+00
urgent	125973.0	0.000111	1.436603e-02	0.0	0.00	0.00	0.00	3.000000e+00
hot	125973.0	0.204409	2.149968e+00	0.0	0.00	0.00	0.00	7.700000e+01
num_failed_logins	125973.0	0.001222	4.523914e-02	0.0	0.00	0.00	0.00	5.000000e+00
logged_in	125973.0	0.395736	4.890101e-01	0.0	0.00	0.00	1.00	1.000000e+00
num_compromised	125973.0	0.279250	2.394204e+01	0.0	0.00	0.00	0.00	7.479000e+03
root_shell	125973.0	0.001342	3.660284e-02	0.0	0.00	0.00	0.00	1.000000e+00
su_attempted	125973.0	0.001103	4.515438e-02	0.0	0.00	0.00	0.00	2.000000e+00
num_root	125973.0	0.302192	2.439962e+01	0.0	0.00	0.00	0.00	7.468000e+03
num_file_creations	125973.0	0.012669	4.839351e-01	0.0	0.00	0.00	0.00	4.300000e+01
num_shells	125973.0	0.000413	2.218113e-02	0.0	0.00	0.00	0.00	2.000000e+00
num_access_files	125973.0	0.004096	9.936956e-02	0.0	0.00	0.00	0.00	9.000000e+00
num_outbound_cmds	125973.0	0.000000	0.000000e+00	0.0	0.00	0.00	0.00	0.000000e+00

2.3 Bereinigung des Datensatzes



2.3 Bereinigung des Datensatzes



Alternative IDS Datensätze

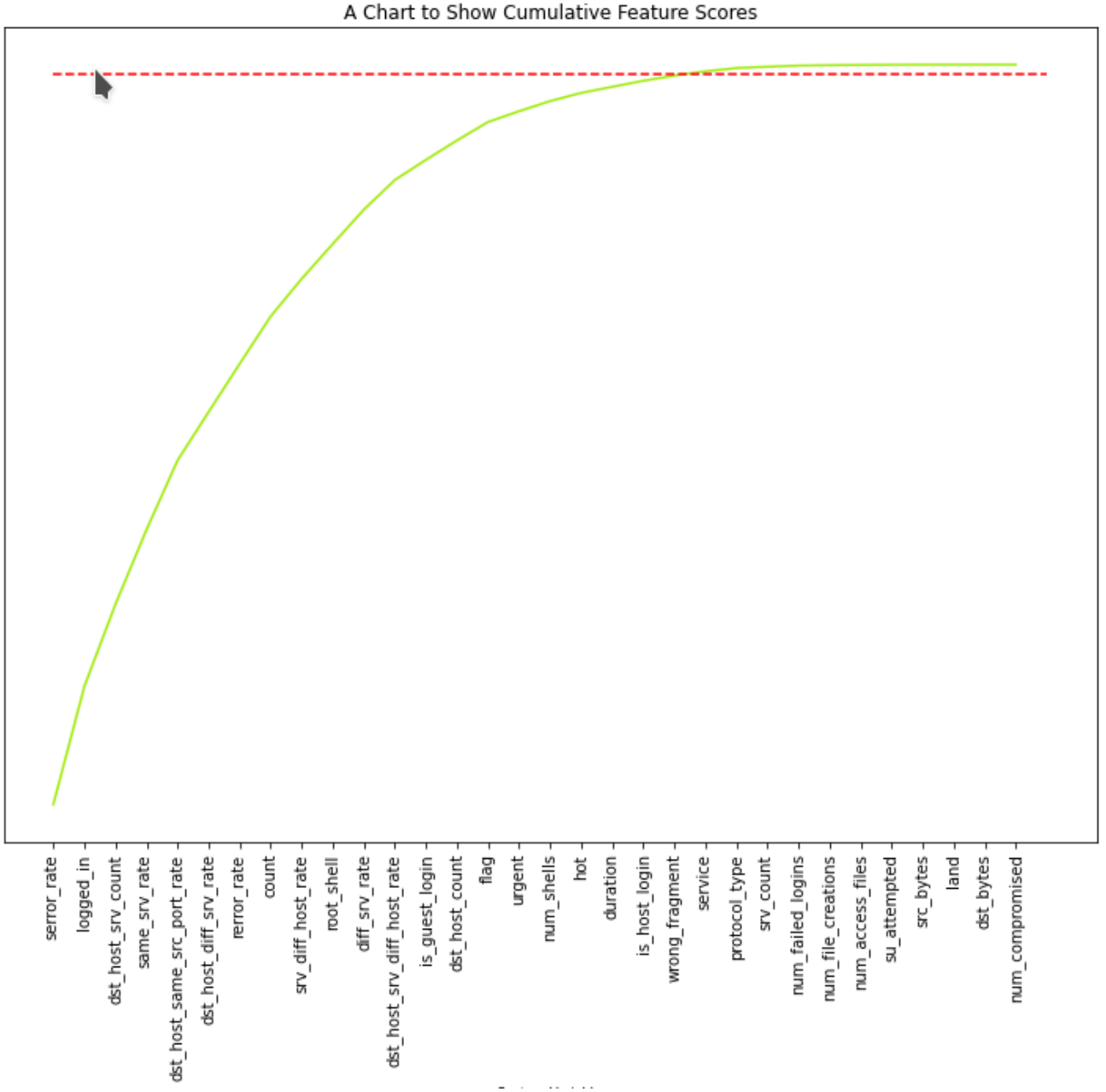
Kyoto2006+
UNSW-NB15
CIC-IDS2017

Roadmap

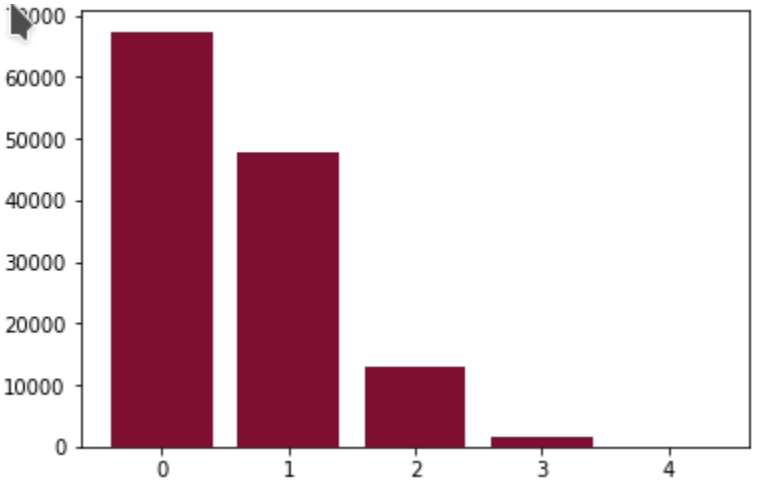
1. Einführung
2. Datensatz
3. Verfolgter Ansatz
    - 3.1 Feature Selection
    - 3.2 Data Imbalance
4. Ergebnisse

Angriffkategorien separat betrachten?

3.1 Feature Selection



3.2 Data imbalance

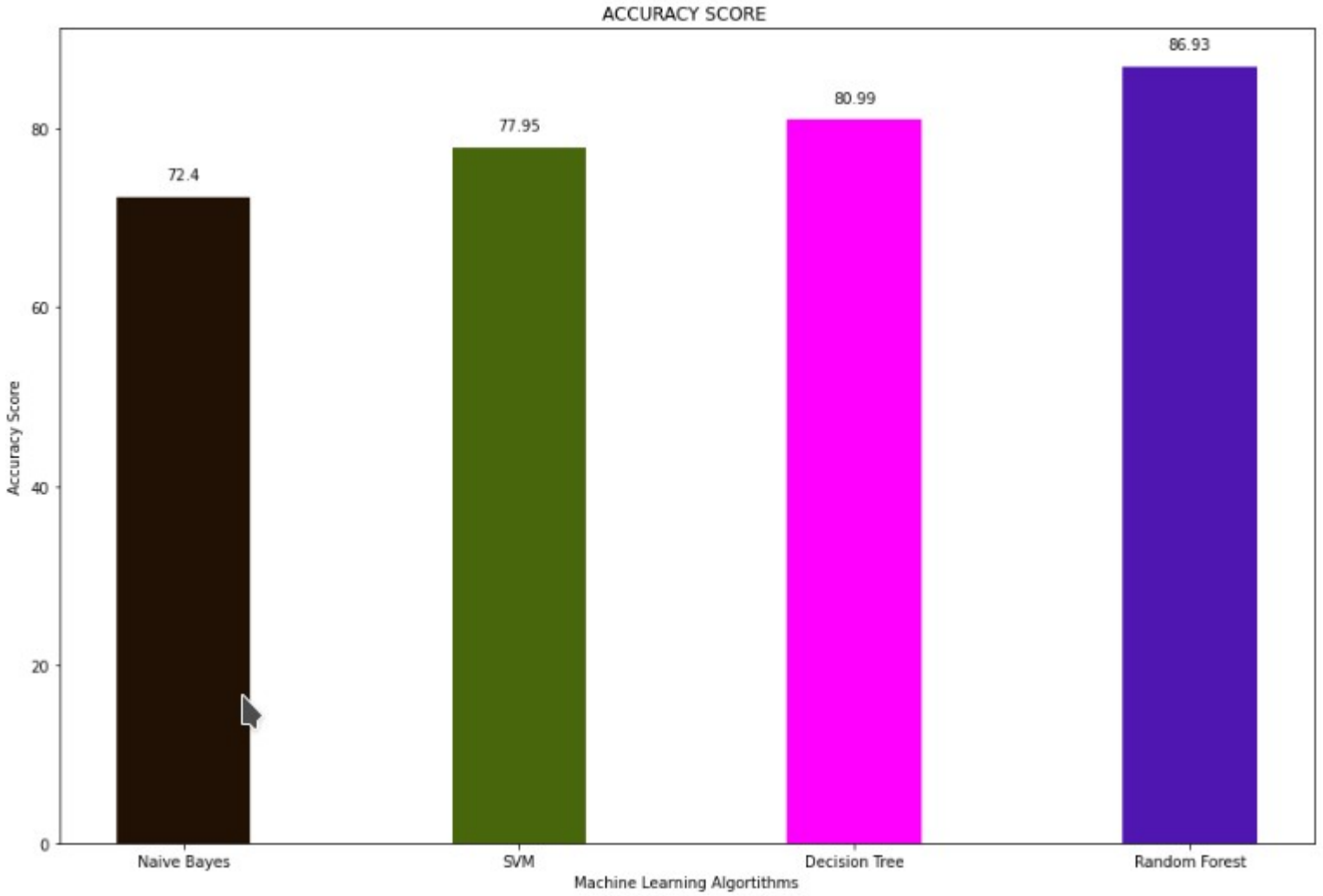


0 = Normal, n=67343 (51.913%)
1 = DoS, n=47646 (36.729%)
2 = Probe, n=12971 (9.999%)
3 = R2L, n=1681 (1.296%)
4 = U2L, n=82 (0.063%)

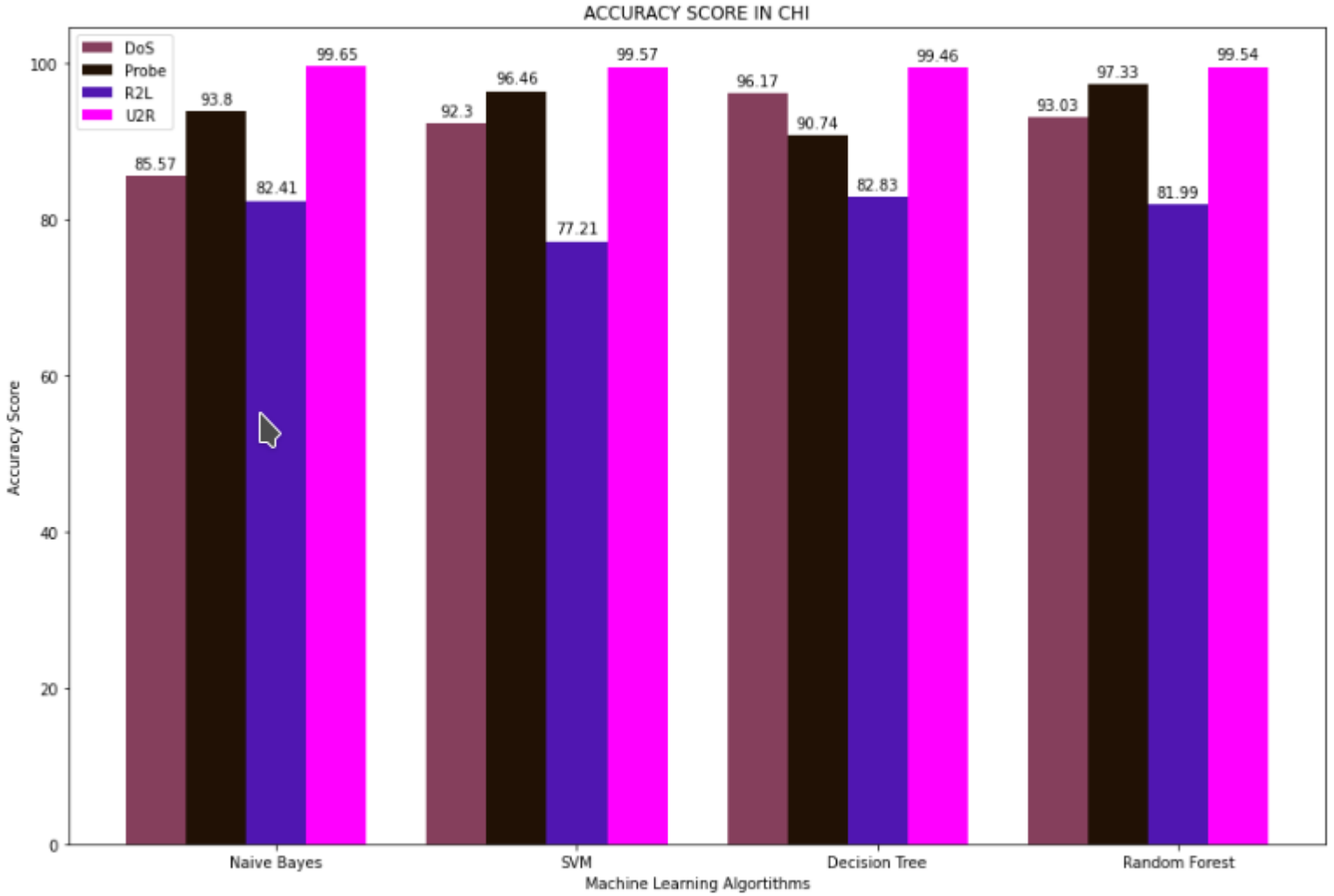
Roadmap

1. Einführung
2. Datensatz
3. Verfolgter Ansatz
   → 4. Ergebnisse

4 Ergebisse


ACCURACY SCORE
Accuracy Score
0
20
40
60
80
72.4
77.95
80.99
86.93
Naive Bayes
SVM
Decision Tree
Random Forest
Machine Learning Algortithms

4 Ergebisse


ACCURACY SCORE IN CHI
DoS
Probe
R2L
U2R
85.57
93.8
82.41
99.65
92.3
96.46
77.21
99.57
96.17
90.74
82.83
99.46
93.03
97.33
81.99
99.54
Accuracy Score
0
20
40
60
80
100
Naive Bayes
SVM
Decision Tree
Random Forest
Machine Learning Algortithms

Vielen Dank für die Aufmerksamkeit!