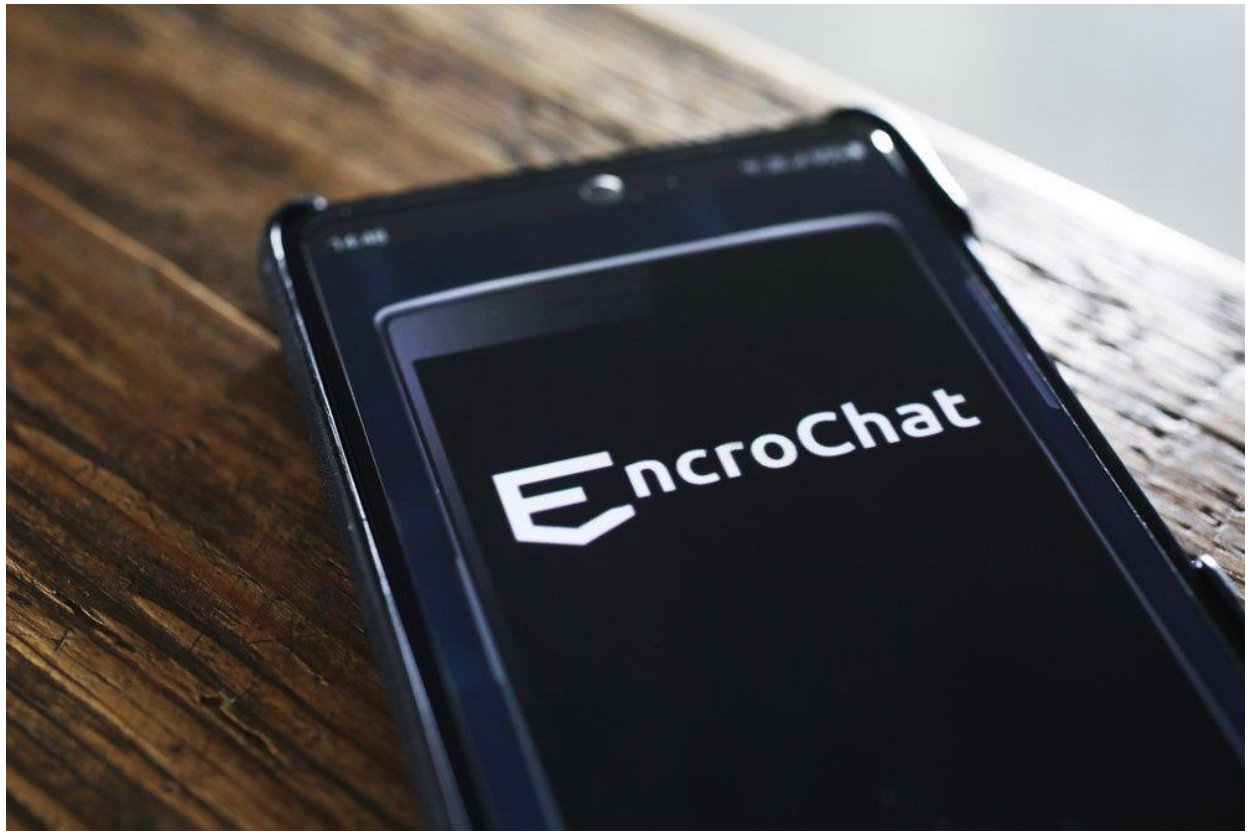
Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid

PaSSw0rdVib3s!: AI-assisted password recognition for digital forensic investigations

DFRWS-EU 2025

Romke van Dijk & Judith van de Wetering

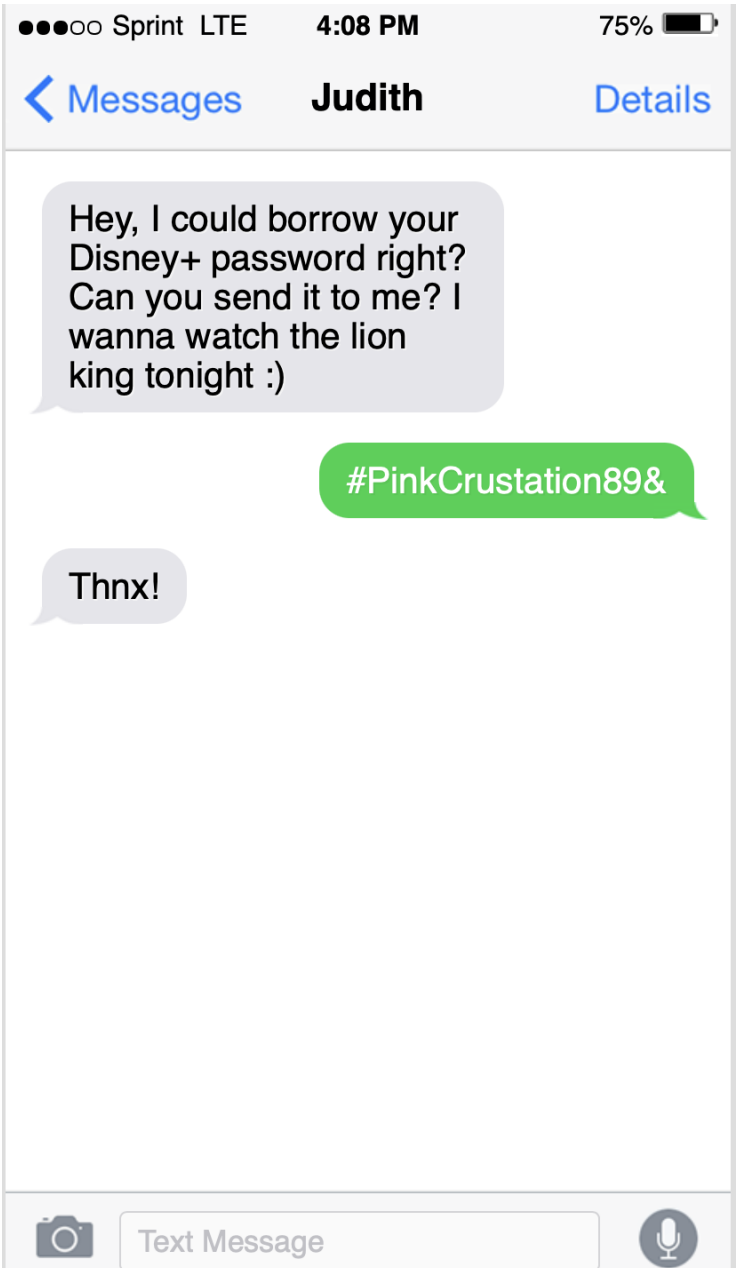
Introduction



Introduction

- People reuse passwords, finding those passwords might result in getting access to other services (Das et al 2014)

Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid




vodafone
9:41 AM
50 %
Chats
Romke
online
Do you have time to work on the presentation for DFRWS later today?
8:45
Yeah sure, I'll hop by your office in a bit.
8:46
Red_Shrimp_23!
9:36
Oh shoot.. please ignore that!
9:36
Message


EncroChat

Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid

today?

zip(components[1:],

Red_Shrimp_23!

presentation

#E29A86;">

Pa$$w0rdVib3s!

Nederlands Forensisch Instituut
Ministerie van Justitie en Veiligheid

Red_Shrimp_23!

Pa$$w0rdVib3s!

#E29A86;">

zip(components[1:],

today?

presentation

Goal

Rank strings, extracted from a mobile device, based on their likelihood of being human generated passwords

AI!


```
Red_Shrimp_23! == password
Pa$$w0rdVib3s! == password
```

```
#E29A86;"></div>     != password
zip(components[1:],  != password
today?               != password
presentation         != password
```

Solutions?

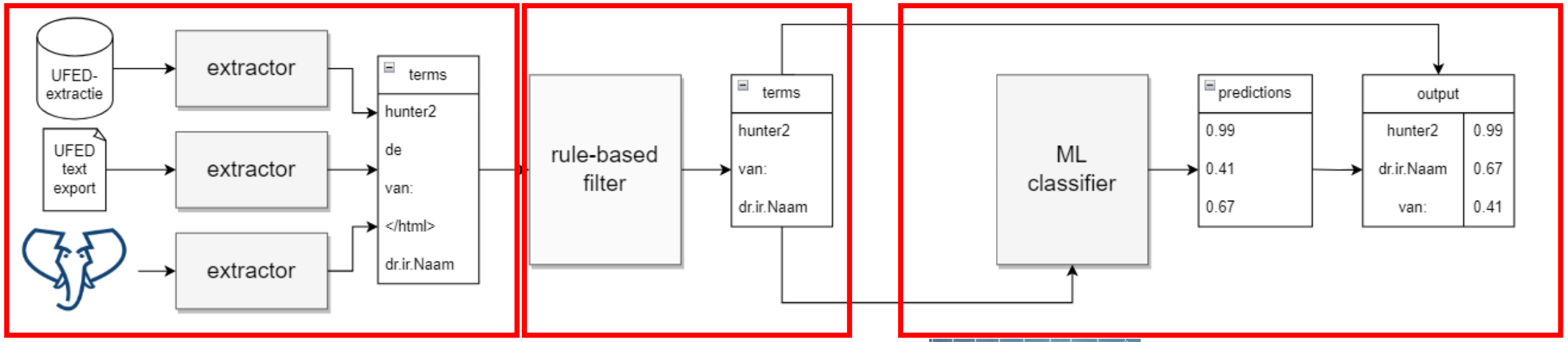
- Previous work:
  - Probabilistic context-free grammar
  - Text Convolutional Neural Networks
    - Focus on Github credential leakage
- Transformers
  - Known to outperform TextCNN

Research Questions

RQ1: What mix of training data yields the best result?

RQ2: Which model performs best?

Method

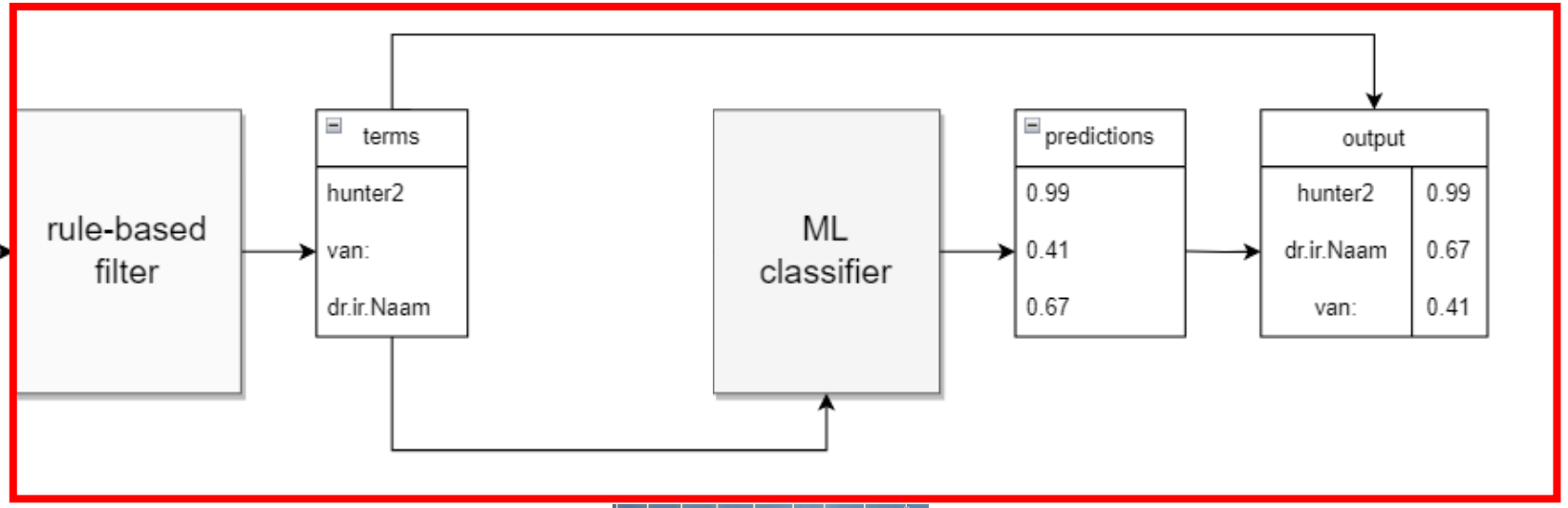


today?
zip(components[1:],
Red_Shrimp_23!
presentation
#E29A86;">
Pa$$w0rdVib3s!
the

today?
zip(components[1:],
Red_Shrimp_23!
presentation
#E29A86;">
Pa$$w0rdVib3s!
~~the~~

```
Red_Shrimp_23!          0.99
Pa$$w0rdVib3s!          0.99
#E29A86;"></div>        0.41
zip(components[1:],     0.31
today?                  0.32
presentation            0.25
```

Method



today?
zip(components[1:],
Red_Shrimp_23!
presentation
#E29A86;">
Pa$$w0rdVib3s!
the

today?
zip(components[1:],
Red_Shrimp_23!
presentation
#E29A86;">
Pa$$w0rdVib3s!
~~the~~

Red_Shrimp_23!	0.99
Pa$$w0rdVib3s!	0.99
#E29A86;">	0.41
zip(components[1:],	0.31
today?	0.32
presentation	0.25

1 Introduction

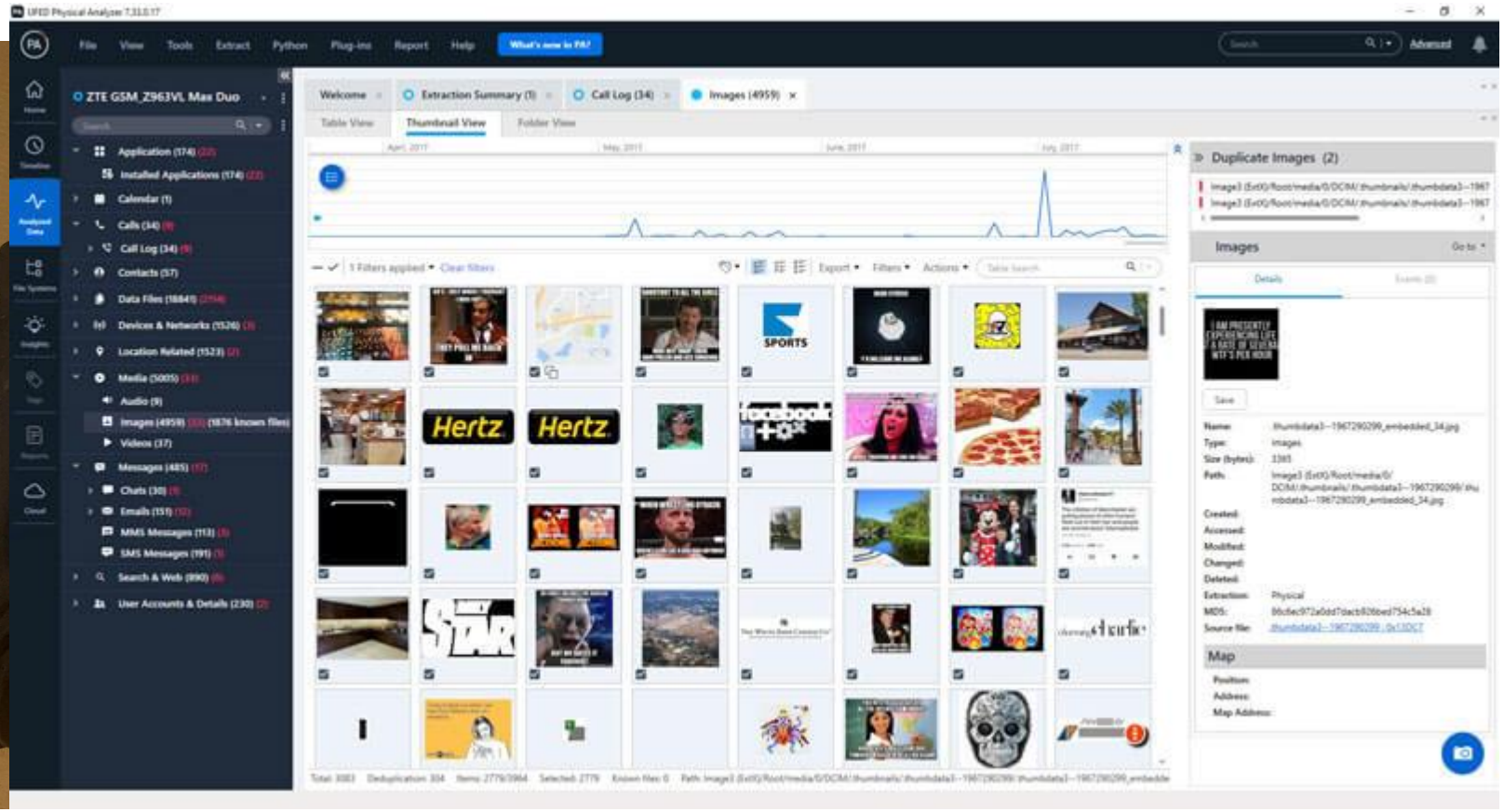
2 Data

3 Models

4 Evaluation

5 Results

UFED


ZTE GSM_Z963VL Max Duo
Welcome
Extraction Summary (1)
Call Log (34)
Images (4959)
Thumbnail View
Duplicate Images (2)
Images
Map
SPORTS
Hertz
Hertz

```
00008101-000C2C383632001E_files_full.zip/private/var/mobile/Containers/Data/Application/B8E90591-EDF8-4A09-B91C
DD:9F:81:37:65:CE
3S<BEL5Ó!Ø-+IND SOS DCS ESC APC ßñ# PAD DC2 RS SUB øÏ PU2 @lÈ' PU2 N« STX ETX SOH NUL SOH STX HOP PAD SUB L½³AxLc{Z SYN ¿úu¨ SO ± DC
3S<BEL5Ó!Ø-+IND SOS DCS ESC APC ßñ# PAD DC2 RS SUB øÏ PU2 @lÈ' PU2 N« STX ETX SOH NUL SOH £ HOP MW 0 HOP PU2 0 US ACK ETX U GS # EOT CAN
com.apple.shortcuts.runtime
D¡ EPA DEL US æ STX ¿ÏÞ DC2 óa SI ä. BEL @ BPH } CSI rí7j[Ý DC1 d DEL SCI »0à SOH ~1 VTS ÈÇ-gçéÕ OSC ûSTW ETB t DEL ÷ètP DC3 RS -óÇ%Íç^IjV'
00008101-000C2C383632001E_files_full.zip/private/var/mobile/Containers/Data/Application/92110031-6BE0-4821-8CD0
com.apple.Home.HomeControlService
Podcasts
C3:A4:25:53:6D:F9
76B22832-F85F-4ECF-B5AA-2C3CB056C288
~P]Z}¨ NBH akL¥?ST íå/¬ËG>î` CAN í?Å¾ STS öxÔ9ÓSå SOH +STX ETX SOH NUL SOH STX HOP SOH 4: SI >+vWÒbÎ%A¦*YCMö{WZ PAD A8Óäòïj PU
4.363222)
com.apple.mobilemail
20G75
(52.048019,
com.apple.ctkui
D4:92:92:D6:F9:99
Û|s NAK ~ SO iL BS i[9 APC g& CSI Ìá¬>Ù IND 5¨ CCH :oÕs¶ GS î08 FF STX ed1200 FF ETX acl1)0 BS FF ETX ock SOH SOH SOH 0 FF EOT odel SOH SOH
48:8B:0A:70:2E:A1
48:8B:0A:70:2E:A0
com.apple.Health.Sleep
8EF3E5FB-20FD-4267-A3EB-323759038C92
E8:C2:2C:81:7F:4F
4.350855)
j ACK EM STS r\#u(ef\®¶ VTS æü´ëÐ_R² STX ;·üpe¹!¸:yÌIñ}^ VTS ¸ STX w PLU NUL GS $TË6Ký$L BS A*ièíß`=,Û¦ùÆvË*4¯ã SS3 SPA qéÞµ ACK
Towers
4509
1513
EOT 0E5¡X£¸[ºÀ> SGCI "Ì RI CSI ¾» ö PM rKÕZãÄÕrãSæû÷ákRf ETB PU2 -jÛ[éêL OSC Ç!Ã CCH wá¸Á FS G.7Ëo'ýk8 DEL ùHµÉ SCI PLU nì½ú STS ñR
CE:6C:42:65:A2:C4
(52.034713,
° HOP ò DLE D¯ SYN j SUB ®£&3 ENQ SUB / DC3 EOT H DC2 ªU~}Fya¯¾¢ SCI }Æh NUL ¿ÎNçìNæ{FÁæ} ETB PU2 ÷Í FS EPA M SO ïú FF CAN 4 PAD ®²C SOH ) D
FA:B8:D8:48:C9:C7
4.338001)
15749
200D7A92-B041-4192-AF9C-27292121D855
4.332366)
```

Normal text file | length : 140.681 lines : 1.948 | Ln : 1 Col : 1 Pos : 1 | Windows (CR LF) | UTF-8 | INS

Training Data

- Passwords
  - Leaked credentials

```
a3f9b47c2d8e6a
f21d8e9b4a6c3d57f
9e4a7b2f18d36c59a2
bfa6d3e9825c47f1d8e9b

Y2hhbGxlbmdlMQ==
c29NZXRoby0xMjM=
U29tZV90ZXh0XyE/QCo
V2lraXBlZGlhIQpOdWxs
```

- Non-Passwords
  - Wordlist English/Dutch
  - Crawls and Chats
  - Carves
  - Encodings

1 Introduction

2 Data

3 Models

4 Evaluation

5 Results

Models

- PCFG
- XGBoost
  - Character N-gram TF/IDF
  - Feature engineering
- Deep Learning (finetuning)
  - PassGPT
  - DistillBERT
  - CodeBERT

1 Introduction

2 Data

3 Models

4 Evaluation

5 Results

Evaluation Data

- Passwords (1000)
  - Datasets
    - RockYou
    - MyHeritage
  - Case data
- Non-Passwords (~250k)
  - UFED Custom Dictionary
    - Apple iPhone 6s plus
    - Apple iPhone 7
    - Apple iPhone 11
    - Huawei P smart
    - Motorola Moto G9 plus
    - Samsung Galaxy J7
    - Samsung S20FE

Evaluation

- Performance of models?
- Presicion@k
  - How many of the top 'k' are passwords?
- Presicion@5: 0.8

Candidate	Score
Sh3llf!shM@ster	0.99
Th!$1sS3cur3	0.97
hunter2$!	0.95
elephant	0.93
Cl@mpocalypse2024	0.91
LetMeIn_456	0.89
CorrectHorse!	0.87
P!nch3r$&Tr@p	0.85
len(info[2])	0.83
Admin123!	0.81
…	…
…	…

1 Introduction

2 Data

3 Models

4 Evaluation

5 Results

Research Questions

RQ1: What mix of training data yields the best results?

RQ2: Which model performs best?

RQ1 – Training data

Chats	Carves	Hex*	feature_xgb precision@100	tfidf_xgb precision@100	PassGPT precision@100
✓	✓	✓	0.36	**0.77**	**0.9**
✓	✓	✗	0.68	0.64	0.83
✓	✗	✗	**0.71**	0.61	0.77
✗	✓	✗	0.18	0.1	0.03
✗	✗	✗	0.16	0.01	0.06

Results of case data

RQ2 - Models

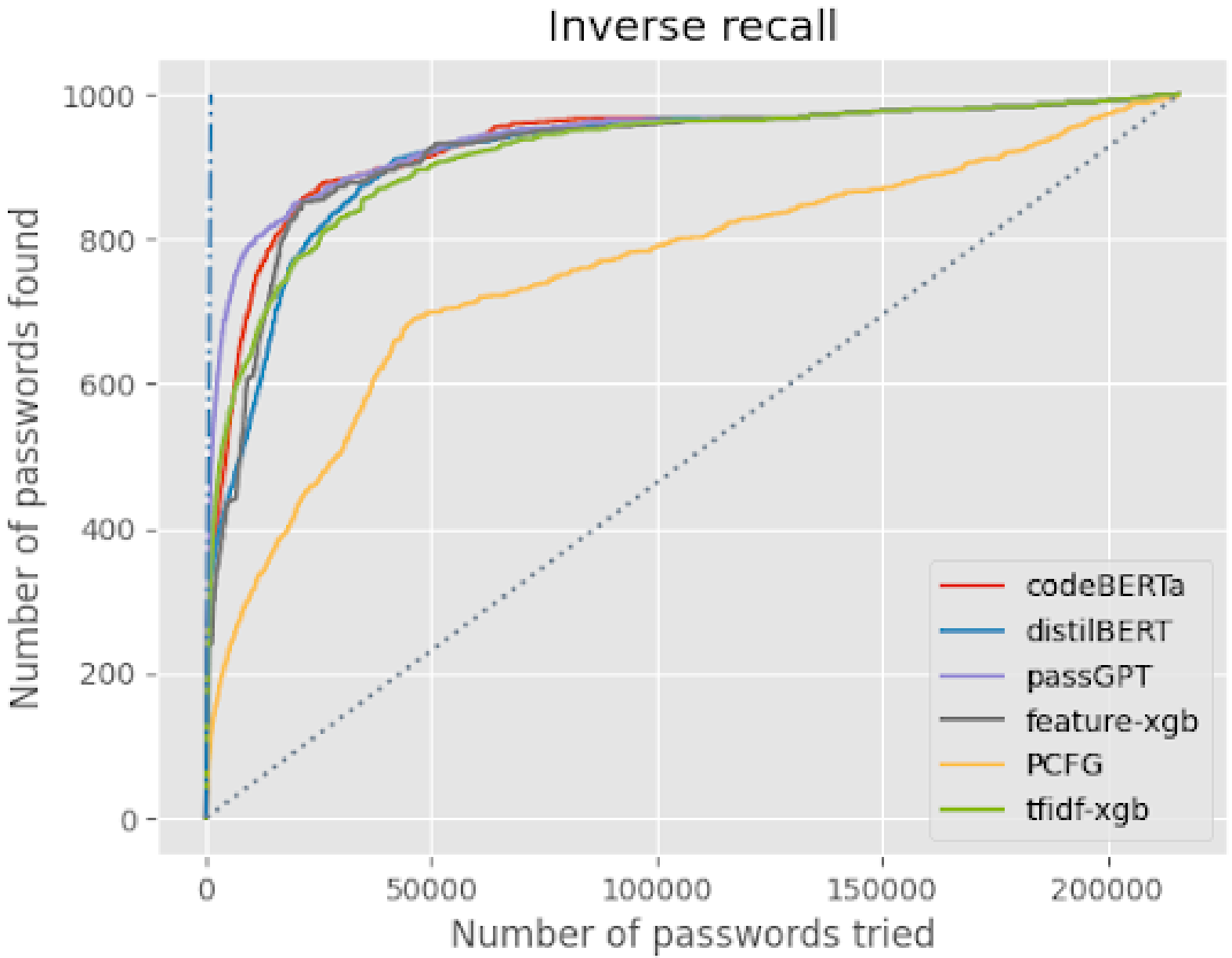
Model	Precision@100	Precision@1000
distilBERT	0.89	0.326
codeBERT	0.83	0.35
passGPT	**0.90**	**0.468**
Feature_xgb	0.36	0.24
Tfidf_xgb	0.77	0.335
PCFG	0.18	0.105

Results of case data

RQ2 - Runtime models

Model	Seconds	Precision@100
distilBERT	21	0.89
codeBERT	20	0.83
passGPT	42	**0.90**
Feature_xgb	**14**	0.36
Tfidf_xgb	38	0.77
PCFG	50	0.18

Results of case data


Inverse recall
Number of passwords found
1000
800
600
400
200
0
0
50000
100000
150000
200000
Number of passwords tried
codeBERTa
distilBERT
passGPT
feature-xgb
PCFG
tfidf-xgb

Conclusion

- Training data important for performance models
- Models outperform existing method
  - PassGPT slower, but more accurate maar nauwkeuriger
  - distilBERT faster, maar but less accurate
  - Usable for case work
- Quality of sorting can be improved

Examples – (MyHeritage)

- Zxcvbnm123 -> 0.9999
  - ✓
- charmed666 -> 0.9999
  - ✓
- renegade13 -> 0.9999
  - ✓
- 62c2e97a024650a9 -> 0.9793
  - ✗
- Greentea! -> 0.1361
  - ✗
- password
- notapassword

Future work

- Improve performance
  - Improve training data
  - Other models
- Context
  - Include location of password
  - What is found 'around' the password?
- Extractors
  - Axiom?
  - Evaluation text extraction UFED

Questions?

Link to paper: