Privalino
Kinder chatten sicher.

WhatsSafe
Die Herausforderung
Warum digitaler Kinderschutz?
51% aller 6-13 Jährigen besitzen ein Telefon

86% davon chatten regelmäßig

Quelle: KIM Studie 2016 (https://www.mpfs.de/studien/kim-studie/2016/)
Cyber-Grooming

Cyber-Mobbing

Radikalisierung
5,3% Erwachsener berichten sexuellen Online-Kontakt mit Minderjährigen.

Quelle: Mikado Studie 2015 (http://www.mikado-studie.de/index.php/101.htm#Online)
5% davon
mit Kindern unter 14

Quelle: Mikado Studie 2015 (http://www.mikado-studie.de/index.php/101.htm#Online)
41% nutzten Täuschungsstrategien
71% baten um Geheimhaltung
35% beeinflussten gezielt mit Geschenken oder Erpressung

Quelle: Mikado Studie 2015 (http://www.mikado-studie.de/index.php/101.htm#Online)
Für diese Kriminellen ist das Internet ein paradiesischer Ort mit direktem und ungestörten Kontakt zu ihren Opfern.

- Julia von Weiler, Innocence in Danger
Zusammengefasst

- Kinder sind im Internet zuhause
- Pädokriminelle Cyber-Groomer nutzen das Internet als Werkzeug zum Missbrauch
- Eltern erfahren nur selten davon
Die Idee

- Live-Analyse von Chats
- Automatisches Erkennen von Gefahren
- Warnung an Eltern und Kind
- Aufklärung über Gefahren und den richtigen Umgang mit ihnen
Wie kriegen wir das hin?
lernt von Erfahrung?

folgt Anweisungen.

lernt von Erfahrung.
Machine Learning

ja/nein?
$1.000 \times 1.000 \times 3 = 3.000.000$
→ 3 Millionen Werte!!
Was ist was?
Deep Learning
Lungenkrebs erkennen

Bildquelle: https://stanfordmlgroup.github.io/projects/chexnet/
Auto fahren

Bildquelle: https://medium.com/swlh/everything-about-self-driving-cars-explained-for-non-engineers-f73997dc60c
This is one of Crichton's best books. The characters of Karen Ross, Peter Elliot, Munro, and Amy are beautifully developed and their interactions are exciting, complex, and fast-paced throughout this impressive novel. And about 99.8 percent of that got lost in the film. Seriously, the screenplay AND the directing were horrendous and clearly done by people who could not fathom what was good about the novel. I can't fault the actors because frankly, they never had a chance to make this turkey live up to Crichton's original work. I know good novels, especially those with a science fiction edge, are hard to bring to the screen in a way that lives up to the original. But this may be the absolute worst disparity in quality between novel and screen adaptation ever. The book is really, really good. The movie is just dreadful.
Aufgabe

Gegeben einen Chatverlauf:
Können wir entscheiden, ob Grooming vorliegt oder nicht?
Eingabe → Algorithmus → Ausgabe
ML Prozess

Daten sammeln → ML Prozess → Daten annotieren

Modell definieren → Modell trainieren → Modell evaluieren
1. Daten sammeln
"Do things that don’t scale"

- Paul Graham, Y Combinator
Erfolg (?)

- Erste Anmache nach 30 Sekunden
- 60 Anmachen innerhalb einer Stunde
- ca. 30.000 Nachrichten insgesamt
Taktiken

- Kanalwechsel (z.B. zu Skype)
- Direkter sexueller Dialog
- Langwierige, durchdachte Manipulation
- Falsche Identitäten
Eingabe

A: bla bla
B: blu blu
A: ....

Algorithmus

Ausgabe
2. Daten annotieren
Grooming-Score von 0 bis 1

Entspricht gefühlter Gefahr zu Zeitpunkt t

Beginnt in jeder Konversation bei 0.5
Eingabe

A: bla bla
B: blu blu
A: ....

Algorithmus

Ausgabe

0.4
3. Modell definieren

ML Prozess

- Daten sammeln
- Daten annotieren
- Modell definieren
- Modell trainieren
- Modell evaluieren
Embeddings → Konv-Block → RNN-Block → Sigmoid-Layer

+ Batch Normalisierung
+ (räumliches) Dropout
Eingabe
A: bla bla
B: blu blu
A: ....
Algorithmus
Ausgabe
0.4
4. Modell trainieren
Training

● Training auf 80% der Daten
● 20% der Daten zum Testen
● Trainingsdauer ~30 Minuten
5. Modell evaluieren

ML Prozess

- Daten sammeln
- Modell definieren
- Modell trainieren
- Daten annotieren
- Modell evaluieren
Zahlen (auf ungesehenen Daten)

- Pearson-Korrelation: 0.96
- Fehler > 0.3: 3%
- Auswertung: < 10 ms
0: na wie gehts dir so hast du skype?

-> (0.882)
hey süße! bist du offen für ältere männer?

wie alt? ich bin erst 13. . .

bin 31 aber voll lieb ;) hattest du schon mal sex?

nein noch nie

-> 0.970
0: hey
1: hii
0: hab bock auf spaß
1: spaß ?
0: intim schreiben
1: oki
0: hab bock auf sex du auch ?
-> (0.972)
hi anna hier ist lea, hast du schon hausaufgaben gemacht?

hi anna! ne noch nicht. was haben wir auf? 😊

buch seite 19, 3 a-e

ah danke!
Zusammengefasst

- Machine Learning lernt von Daten
- Deep Learning revolutioniert seit 2010 alle denkbaren Bereiche
- Erkennt Grooming in Chats
Und in der Praxis?
Privalino
Kinder chatten sicher.

WhatsSafe
Hey, anscheinend versucht gerade jemand, mit dir über etwas sehr Privates zu sprechen:

**Bist du sicher, dass du das willst?**

Denke mal darüber nach:
- Kennst du diese Person wirklich?
- Was einmal im Netz ist, bleibt im Netz!
Mit neuesten Technologien: Richtige Aufklärung im richtigen Moment!
Diskussion
Kontakt

Patrick Schneider
Geschäftsführung
patrick@privalino.de

Max Schumacher
Technische Leitung
max@privalino.de