

## Einsendeaufgabe 1

### EXERCISE DSI-01

*Reserve at least 30 min and read the latest "State of AI Report!" (Benaich / Hogarth)*

*Write me 3 paragraphs of 3 points you found interesting or got stuck in your mind.*

Folie 22:

Die Aspekte zum Erlernen von Sprachmodelle und Sequenzen fand ich sehr interessant, konnte mir aber kein Praxisbeispiel für Sequenzen in Aminosäuren Ketten vorstellen. Daher war die Folie zur Sprache der Infektionskrankheit Covid 19 sehr aufschlussreich. Mir war mir nicht bewusst, dass Data Science auch in den Bereichen Gesundheit und Forschung entsprechend eingesetzt wird. Und damit einen sehr konkreten Nutzen zum Schutze der Gesundheit der Bevölkerung hat.

Folie 25:

Auch der Einsatz von Data Science im Bereich der Medikamentenforschung war mir bisher unbekannt. Durch eine experimentelle Erprobung von neuen Medikamenten und ihren Dosen kann die Sicherheit von Medikamententest an Menschen erheblich verbessert werden. Ebenso kann so schon frühzeitig abgeschätzt werden welches Medikament in welcher Dosis eingesetzt werden kann, um eine gewünschte Wirkung zu erzielen.

Folie 33:

Der Aspekt, dass ein Vorgehen auch rechnerisch sinnvoll und im Verhältnis zum Nutzen stehen muss, ist sehr interessant. Im konkreten Fall von Video-Beschreibungen nutzt ClipBERT nur wenige kurze Ausschnitte, um ein aussagekräftiges Ergebnis zu erreichen. Würde das ganze Video durch die Software analysiert werden wäre der rechnerische Aufwand nicht im Verhältnis zum Ergebnis.

### EXERCISE DSI-02

*Make up your own example of a machine learning task with a binary classification with your own numbers. Like e.g. recognizing skin cancer. Calculate precision and recall.*

Das von mir gewählte Beispiel ist ein Schwangerschaft Test

1. Person ist schwanger + Test sagt Person ist schwanger → richtiger positiv **tp**
2. Person ist schwanger + Test sagt Person ist nicht schwanger → falsch negativ **fn**
3. Person ist nicht schwanger + Test sagt Person ist schwanger → falsch positiv **fp**
4. Person ist nicht schwanger + Test sagt Person ist nicht schwanger → richtig negativ **tn**

Ein Beispiel hierzu wäre:

15 Schwangerschaft Tests werden durchgeführt, davon geben 10 ein positives Ergebnis an und in 8 Fällen liegt wirklich eine Schwangerschaft vor.

### Precision

In dem gewählten Beispiel liegt die Precision bei:

$8/(8+2)=0,8$  entspricht einer Precision von 80%

### Recall

$8/(8+7)=0,53$  entspricht einem Recall von 53%

### EXERCISE DSI-03

*What is the F1 measure of your example? (real calculation needed here).*

$F1 = 2 * (\text{precision} * \text{recall} / (\text{precision} + \text{recall})) = 0,64$  entspricht einem F1 Wert von 64%