

## 1 Lernziel 1

Lernziel 1 war das Erklären und Bewerten einer historischen und einer aktuellen Entwicklung der AI. Diese Erfüllung dieser wird im Folgenden einzeln Erklärt:

### 1.1 Historische Entwicklung - NLP

Für die historische Entwicklung der AI, habe ich mich mit NLP-Methoden außerhalb von Machine- und Deep-Learning auseinander gesetzt. Die alten Methoden, Hidden-Markov und TF-IDF sind zwar deutlich *simpler* als aktuelle Entwicklungen im NLP Bereich wie Transformer-Modelle und Co., aber auch diese *alten* Verfahren finden immer noch Anwendung. Es war interessant sich nicht nur die Funktionsweise dieser Verfahren anzusehen, die teilweise bereits bekannt war, sondern zu sehen welchen Einfluss sie auch heute noch haben, weil sie für ihre Anwendungszwecke vollkommen ausreichend sind oder bessere Alternativen einfach zu entwicklungs- oder rechenintensiv sind. Oder aber, dass neuere Verfahren auf diesen alten (schon in den 50er Jahren verwendeten) Methoden aufbauen, oder sie inhärent nutzen. Praktisch wurde TF-IDF selbst auch schon im letzten Semester eingesetzt, um Reddit in bestimmten Kategorien auf aktuelle Trends zu untersuchen: [https://github.com/FatManWalking/Reddit\\_Trend](https://github.com/FatManWalking/Reddit_Trend), damals mit Florian Köhler gemeinsam. Zugehörige

Einreichung: [https://github.com/FatManWalking/learning\\_ai/blob/main/History\\_of\\_NLP.pdf](https://github.com/FatManWalking/learning_ai/blob/main/History_of_NLP.pdf)

### 1.2 Aktuelle Entwicklung - Adversarial Attacks

Für die aktuelle Entwicklung der AI, habe ich mich mit den Adversarial Attacks befasst. Dies ist ein Thema das erst circa in den letzten 5-6 Jahren wirklich groß geworden ist und mit dem ich mich umfassend in meiner Bachelorarbeit beschäftigt habe. Die Erkenntnisse, die ich in meiner Bachelorarbeit gewonnen habe, noch Revue passieren zu lassen und in wenigen Seiten zusammenfassen zu müssen, hat mir die Möglichkeit geben mich nochmal mit diesem Thema auseinander setzen zu können. Hier nimmt die Motivation hinter einem solchen Angriff und auch hinter der Verteidigung eine deutlich zentralere Rolle ein, weil es nicht möglich ist in einem Artikel dieser Länge adäquat auf technische oder mathematische Details eingehen zu können. Und auch die Anzahl der Methoden ist so umfassend, dass man sie kaum alle auflisten kann. So habe ich beim Verfassen des Artikels erst von CopycatCNN erfahren: Ein Verfahren das ich persönlich sehr spannend finde und für die Bonusabgabe in Machine Learning ausprobiert habe ([https://github.com/FatManWalking/ml\\_bonus](https://github.com/FatManWalking/ml_bonus)), auf dem KI-Portfolio war schließlich kein Platz mehr dafür.

Zugehörige Einreichung:

[https://github.com/FatManWalking/learning\\_ai/blob/main/Adversarial\\_Attack.pdf](https://github.com/FatManWalking/learning_ai/blob/main/Adversarial_Attack.pdf)

## 2 Lernziel 2

Lernziel 2 war der Vergleich und die kritische Bewertung zweier gesellschaftlicher Auswirkungen von AI-Entwicklungen. Hierzu wurde ein Podcast mit dem Namen "Lust auf AI" gemeinsam mit Klemens Gerber eingereicht. Ziel dieses Podcasts ist es die Auswirkungen vergangener und zukünftiger Entwicklungen in der Forschung zu AI auf das wirtschaftspolitische und sozioökonomische Klima der Vergangenheit und der Zukunft zu thematisieren. Wir sind dabei so vorgegangen, dass wir uns auf die Themen "KI und die Arbeit" und "der Staat und KI" einigten und dann in einer ersten Phase eigenständig zu beiden Themen recherchierten. Aufgrund unserer unterschiedlichen sozialen und politischen Ansichten, sowie der unterschiedlichen Literatur, die wir gelesen hatten, konnten wir mit einander über die Thematik diskutieren und ein grobes Skript für den Podcast aufbauen, in dem beide ihre Ansichten ausreichend betrachtet sahen. Dadurch konnte sich nicht nur ein individuelles Bild dieser Sachverhalte gemacht werden, sondern auch mit jemand anderem, mit anderen Ansichten zur Thematik, auch außerhalb des Podcasts diskutiert werden. Dies hat uns, denke ich, beiden geholfen nicht nur den Einfluss von KI besser zu verstehen, sondern auch zu merken, wie unterschiedlich Menschen, selbst mit ähnlichen Interessen und quasi aus dem Fachbereich, zu dieser Thematik stehen. Aufgrund der Länge und des thematischen Umfangs ist der Podcast als Inhalt konzipiert, der gleich zwei

Portfolio-Abgaben abdecken soll. Weiterhin wurde in der Erstellung des Portfolios der Fokus auf dieses Lernziel gelegt, weshalb eine entsprechende Gewichtung bei der Bewertung des selben von mir ausdrücklich gewünscht ist. Die Themen des Podcasts decken die Anforderungen bezüglich der Beschäftigung mit gesellschaftlichen Auswirkung der AI-Entwicklung ab und sollen das zweite Lernziel somit vollständig erfüllen. Zugehörige

Einreichung: [https://github.com/FatManWalking/learning\\_ai/blob/main/Lust%20auf%20AI.mp3](https://github.com/FatManWalking/learning_ai/blob/main/Lust%20auf%20AI.mp3)

### 3 Lernziel 3

Für Lernziel 3: Das Einreichen eines regelbasierten und eines genetischen Verfahrens habe ich mich um eine kombinatorische Umsetzung innerhalb des alten Atari-Klassikers „Pong!“ gekümmert. Der regelbasierte Algorithmus ist hierbei so simpel, wie effektiv: Über eine einfache Imitierung der Position des Balles, über eine Folge von if-Bedingungen ist die Plattform bei gleicher Geschwindigkeit jeder Zeit in der Lage den Ball zurück zu spielen. Ein aufwendigeres Verfahren wäre in diesem Kontext überflüssig. Anders sieht es aus, wenn der Ball schneller ist als die Plattform selbst und somit dem Ball nicht eins-zu-eins gefolgt werden kann. Diesen Fall ist der regelbasierte Algorithmus zwar theoretisch über Berechnungen in der Lage, doch die fehlerfrei Implementierung wäre deutlich aufwändiger. Deshalb wird der regelbasierte Algorithmus in einer zweiten Instanz dazu genutzt Generationen von genetischen Pong-Gegenspielern zu trainieren/entwickeln. Hierzu habe ich einmal versucht ein Netz über Screencapturing des Spiels per visuellen Input zu trainieren, dieses Verfahren ist leider fehlgeschlagen, da die Prozesse mit einander interagieren mussten, ich mich nicht in asynchrone Programmierung eingearbeitet habe und die Prozesse sich entsprechend blockiert haben. In einem zweiten Versuch habe ich ein Feedforward-Netzwerk nur basierend auf den X- und Y-Koordinaten für alle Objekte im Spiel trainiert. Dies funktionierte überraschend gut in einzelnen Instanzen. Zusätzlich habe ich ihm Rahmen des Machine-Learning-Projekts Duel-Q-Learning erfolgreich für Doom ([https://github.com/FatManWalking/the\\_fat\\_bot](https://github.com/FatManWalking/the_fat_bot)) implementiert und fühle mich entsprechend sicher was Reinforcement-Learning-Algorithmen angeht. Was halb ich hier die „etwas einfachere“ Variante mit einem genetischen Algorithmus gewählt habe. Zugehörige Einreichung: [https://github.com/FatManWalking/learning\\_ai/tree/main/Pong-DQN](https://github.com/FatManWalking/learning_ai/tree/main/Pong-DQN)

### 4 Lernziel 4

Lernziel 4 war das Erklären und Anwenden eines generativen Verfahrens jeglicher Art. Hierzu entschied ich mich im Rahmen einer der aufgetragenen Hausaufgaben an einem Variational-Autoencoder (VAE) für die Generierung neuer Pixelcharaktere zu setzen. Hierzu kam mir die Idee, dass ,da der Datensatz hier selbst nicht all zu groß war, als Encoder eines der bekannten vortrainierten Netze für Image Recognition zu nehmen (VGG11), auch wenn die Ergebnisse mit einem simplen Encoder aus drei linear Layern auch bereits recht annehmbar waren. Die Ergebnisse dabei waren eher schlecht, da das Netz deutlich zu komplex für den Anwendungszweck geworden war. Ein deutlicher simplerer Encoder basierend auf drei Linear-Layern war hier deutlich besser im Encoder-Decoder-Prozess. Hierbei war zu sehen, dass das Netz dabei auch gut generalisierte und beispielsweise bei einem Character mit einem Schild, diesen weg lies und stattdessen eine normale Hand an dieser Stelle zeichnete. Ebenfalls überraschend war deshalb, dass die neu generierten Bilder eher der Pixelnoise entsprachen aus der sie entstehen sollten, als den gewünschten Charakteren. Zuletzt wurde ein Netz mit einigen Convolution- und Linearlayern verwendet. Dies funktionierte in unterschiedlichen Hyperparameter-Kombinationen wirklich gut, sowohl im Encoder-Decoder als auch bei der Neugenerierung.

Zugehörige Einreichung: [https://github.com/FatManWalking/learning\\_ai/tree/main/VAE](https://github.com/FatManWalking/learning_ai/tree/main/VAE)