A3

Genau. Ich nämlich auch. Wir machen immer natürlich ein Backup. Und dann würde ich auch schon anfangen mit der ersten Interviewfrage. Naja, stellen Sie sich oder stell du dich in dem Fall doch mal kurz vor.

I9

Natürlich gerne. Also, wie Sie schon wissen, XXX XXXXXXXXX ist mein Name. Ich komme aus dem Iran. Seit fünf Jahren wohne ich hier in XXXXXX. Erst als Student im Bereich Angewandte Informatik, aber mit Schwerpunkten Richtung Machine Learning und künstliche Intelligenz. Und im September habe ich mein Studium abgeschlossen und seit Januar arbeite ich als Softwareentwickler bei einem Unternehmen hier in XXXXXX. Und ich kann sagen, die Aufgaben sind hauptsächlich Softwareentwicklung, aber dabei ab und zu bekomme ich die kleine Aufgabe im Bereich von Machine Learning, also die Anwendung von Machine Learning in diesem Unternehmen.

A3

Oh, perfekt. Das macht sie zu einem sehr geeigneten Teilnehmer in dieser Studie. Dann, was genau ist dein höchster Bildungsabschluss? Also du hast gemeint, dass du studiert hast.

I9

Master

A3

Deinen Master hast du auch gemacht. Okay. In welchem Unternehmen arbeitest du denn? Was ist so das Geschäftsmodell?

I9

Das Unternehmen heißt XXXXXXXXXXXX. Hauptsächlich ist es ein Unternehmen, das sich mit dem XXXXXXXXXXXXX in Deutschland beschäftigt. Aber es gibt auch einen anderen Service, den dieses Unternehmen anbietet. Und das ist die Haus [schlechte Verbindung]. Das ist natürlich die [schlechte Verbindung] und Anforderung von diesem Unternehmen. Also es gibt eine EDV-Abteilung. Und die EDV-Abteilung bereitet alle diese Software und digitale Lösungen für diese Unternehmen vor.

A3

Kannst du nochmal wiederholen, was vor der EDV-Lösung war? Da bist du nämlich kurz abgehakt worden.

I9

Also hauptsächlich geht es um die Entwicklung der Software und der Website des Unternehmens. Nicht nur Entwicklung, sondern Weiterentwicklung und Maintenance, so würde ich sagen. Aber ab und zu kommt auch eine andere Anforderung, wie zum Beispiel Fraud Detection bei Kunden.

A3

Okay, alles klar. Gut, dann beschreib doch mal kurz deine Rolle wirklich genau in dem Unternehmen und ja, was machst du genau?

I9

Ich bin Softwareentwickler im Bereich Datnet und gerade habe ich eine Aufgabe, um eine interne Software für die Callcenter zu entwickeln.

A3

und hast du benutzt du dafür KI zufällig zufälligerweise?

I9

In diesem Anforderungsbereich nein, aber andere Aufgaben wie zum Beispiel, wie ich gesagt habe, Fraud Detection, ich nutze, ja, ich bin gerade dabei, um erstmal die Daten irgendwie herauszuziehen und Data Processing und also die statistische Analyse durchzuführen und Klassifizierung. Also, kann es gehen.

A3

Okay, Klassifizierung. Ist das auch die Art, wie du hauptsächlich mit KI schon gearbeitet hast, dass du eben eine Klassifizierung umgesetzt hast mit einem KI-Modell oder wie hast du noch damit gearbeitet?

I9

Ob ich schon früher mit Machine Learning Verfahren gearbeitet habe, so meinen Sie?

A3

Ja, das auch, ja.

I9

Bei meiner Abschlussarbeit, ich habe auch bei einem anderen Unternehmen hier in XXXXXX gearbeitet, das heißt XXXXXX XXXX. Es ist ein Unternehmen, das digitalisiert das Projektmanagement Vorgehensweise für die Baufirmen. Und die hatten ein Konzept und dafür musste ich auch die Image-Clustering-Modelle benutzen, um quasi ein Label Dataset für die Image-Classification-Modelle wie zum Beispiel YOLO Version 8 oder die anderen Modelle zusammenzupacken.

A3

Okay, alles klar. Wie arbeitest du, wenn wir jetzt nochmal auf dein aktuelles Unternehmen zurückkommen, zum Beispiel auf den Fraud Detection Fall, wie arbeitest du mit Machine Learning Modellen oder mit KI im Allgemeinen?

I9

Wie arbeite ich? Kannst du die Frage ein bisschen mehr erklären?

A2

Ich kann, glaube ich, auch sagen. Also baust du das System, baust du alles from scratch, machst du alles selbst oder suchst du dir irgendwelche Packages von GitHub oder wie? Also ich kann mir vor allen Dingen überhaupt nicht vorstellen, wie sowas funktioniert.

I9

Ach so, ich kann sagen, also der größte Schritt beim Machine Learning und Training und sowas ist erstmal die Data sammeln und dann nach der Datasammlung muss man die Feature Engineering und Data Processing durchgehen

A2

Und das machst du selbst?

I9

Das mache ich selbst. Im Moment, das mache ich selbst, weil, ja, und dann kommt die Modellselection-Phase und welche Klassifizierung-Modell oder Clustering-Modell(?) kann geeignet für diese Anforderungen sein, ist quasi eine andere, sorry, Analyse-Phase. Es gibt ja verschiedene, so wie wir es wissen, zahlreiche Klassifizierung- und Clustering-Models und ich muss dann dabei einige von denen aussuchen und testen und die beste, als die Ergebnisse zurückgeben. Aber keine, im Moment, [Schlechte Verbindung]

A3

so den letzten Teil hat man jetzt nicht so ganz verstanden

I9

Hallo?

A3

jetzt bist du wieder flüssig vielleicht noch mal versuchen

I9

Also die Library... Also die Thingworks, die ich für die Machine Learning nutze, ist hauptsächlich CycleLearn(?) und Tensorflow(?). Die nutze ich im Moment. Können Sie mich gut hören?

A3

Ja, so halbwegs. Vielleicht würde es Sinn ergeben, vielleicht auf mobile Daten umzuwählen. Das hilft oft tatsächlich. Ansonsten hin und wieder geht es gut, aber manchmal hackt es komplett. Da können wir leider nichts verstehen.

I9

die wifi ist signalisch in ordnung ich glaube sie hören mich jetzt besser

A3

Ja, jetzt gerade geht es auf jeden Fall.

I9

Hören Sie mich jetzt gut?

A3

Ja, ja, okay, gut, dann versuchen wir es mal weiter. Perfekt. Du hast es ja schon ganz eindrücklich beschrieben, du hast ein sehr breites Aufgabenspektrum, also du machst ja sowohl Data Preparation als auch Modelauswahl und dann Deployment. Was sind denn so die größten Herausforderungen in diesem Deployment-Prozess? Jetzt ist er weg.

I9

in diesem Deployment Prozess ist die also die große Herausforderung ist natürlich wir haben gar keine sorry

A3

ich dachte dass du ja alles gut

I9

die große Herausforderung ist für die dataset selbst in gerade wir haben keine Labels in unserem Datenbank alle daten die wir kriegen sind einfache Datensätze aus den Datenbanken ohne Klassifizierung wir können ja nicht sagen ob diese Person also es geht um XXXXXXXXXXXXXXX XXXXX XX XXX XXXXXXXXX XXXX XXXXXX XXXX XXX XXXXX XXXXX XXXX XXX XXXXXXX XXXXX XXX XXXX XXXXXXX XXXXXX XXXXX XXX XXXXXXXXXX XXXXXX XXX XXXXX XXXXXXXXXX XXX XXXX XXXX XXX XXXXXX XXX XXXXX XXXXX XXXXXXXXXXXXXXX XXX XXX XXX XXXXXXXXX XXX XXXXXX XXXX XXX XXXXX XX XXXXXX XXX XXXX XXXXX XX XX XXXXX XXXXX XXXXXX XXX XXX XXXXXXXX XXXX XXX XXX XXXXXXXXXXXX XXXXXXXX deswegen die Idee die ich habe ist einfaches Label quasi Dataset und ich muss irgendwie eine Art und Weise nachdenken und irgendwie diese Millionen Datensätze die ich habe eine Label erstellen und deswegen habe ich mir erstmal Cluster Models also unsupervised Models vorgestellt und dann vielleicht durch diese Clustering Models kann ich bessere Einblick in daten haben und irgendwie die auf also den [Anfang von labeling reduzieren und dann kann ich klassifizieren und nutzen] (undeutlich)

A2

Ich muss jetzt nochmal nachfragen, weil ich vieles einfach nicht so schnell verstehe wie XXXX. Ihr seid ja die beiden Experten. Du sagst Clustermodelle… Wie kann ich mir das vorstellen? Du musst ja ein Clustermodell auswählen. Wie sieht so ein Entscheidungsprozess bei dir aus? Also nimmst du das Beste, wie findest du raus, welches jetzt das Beste für dich ist?

I9

Das würde ich aus den Testergebnissen sagen. Also ich kann ja nicht so einfach sagen, welches Cluster im Modell die beste für meine Anforderung ist. Ich muss natürlich an den Kriterien hier nachdenken und basierend darauf kann ich sagen, welche Klasse im Modell funktioniert hier am besten. In dieser Phase bin ich noch nicht herangekommen, aber das muss ich natürlich auch durchdenken.

A2

Und dann so Clustermodelle, wo finde ich die denn? Sorry, ich kann mir das wirklich nicht vorstellen.

I9

wo finde ich diese klasse Models also…

A2

Oder entwickelst du die selber?

I9

Also die Modelle sind schon bereit. Also die klassischen Clustering-Modelle, die meisten sind durch die APIs von Scikit-Learn-Bibliotheken vorhanden. Andere, wie zum Beispiel die neuronalen Netzwerke basierte Clustering-Modelle, die sind auch teilweise vorhanden in TensorFlow, Library, aber nicht so viel. Die neuesten Modelle recherchiere ich selber und nutze ich deswegen manchmal die GitHub-Codes oder die Quellen, die im Internet sind, um die Modelle selber nochmal zu programmieren und zu erstellen.

A2

Ich verstehe. Das heißt, du gehst wirklich in die Library und dann guckst du dir halt die Parameter an und entscheidest danach, welches du ausprobierst?

I9

Besser zu sagen, also für die Entscheidung, welches Modell das beste Modell ist, man kann Testerror und diese Fehlerrate von dem Modell miteinander vergleichen. Aber welches Modell suche ich aus, aus tausenden? Das ist nicht einfach zu sagen. Ich würde dann sagen, also je mehr meine Arbeitszeit mir erlaubt, je mehr diese Anforderungen geduldig sind, ich würde dann desto mehr Clustermodelle berücksichtigen und testen. Natürlich kann man nicht alle tausende Modelle hier testen, aber vielleicht nach Recherchierungen kann ich sagen, okay, diese Anforderungen, also diese Problemstellungen, die ich im Internet gefunden habe, ist ähnlich wie meine und das hat diese Klasse Modell X als die beste berichtet. Und deswegen suche ich das aus und teste ich auf meinen Anforderungen und versuche herauszufinden, ob das mehr passt.

A2

Cool. Ja, jetzt, ich glaube, ich habe es verstanden. Danke.

A3

Okay, super. Du hast ja gerade über Clustering-Modelle quasi in diesem Data Preparation, Data Processing Step geredet. Wenn wir jetzt nochmal weitergehen und du an die Klassifikation gehst, zum Beispiel, wenn du dich zwischen verschiedenen neuronalen Netzen entscheiden musst, wie entscheidest du dich da? Unterscheidet sich das vielleicht auch so ein bisschen zwischen den Clustering-Verfahren und den neuronalen Netzen?

I9

Sorry das habe ich noch nicht so völlig verstanden ein Clustering modell kann auch als Neronales Netzwerk bestehen

A3

Also ich kenne halt diese Unterscheidung, dass Clustering eher unsupervised learning ist und dass neuronale Netze ja supervised learning sind.

I9

Es gibt doch neuronale Nezwerke, die unsupervised auch sind. Es gibt viele Cluster Modelle, die aus Autoencoders bestehen oder aus convolutional neural networks bestehen und vieles.

A3

Okay, alles klar. Gut, danke für die Erklärung. Gibt es denn so Hauptkriterien, du hast es gerade schon genannt, auf die du achten würdest nochmal, also Benchmarks in der Entscheidung?

I9

Benchmark, das muss ich bei der Aufgabe, die ich habe, noch mehr tiefer berücksichtigen. Welche Clustering-Modelle kann mir die besten Ergebnisse ermitteln? Im Moment kann ich unter die Clustering-Modelle nicht entscheiden. Ich habe eine Design-Matrix, ich gebe es dem Clustering-Modell A, B, C und Clustering-Modell A ermittelt sich der Test-Error, sagen wir mal 80%, Clustering-B 90%, die andere 95%. Aber basiert auf diesen Test-Errors, man kann immer noch nicht entscheiden, weil die Clustering-Modelle, die Clusters, kann immer noch fehlerhaft sein. Was ich am Ende entscheiden kann, ist, wenn ich aus diesen Clusters selber durchgehe und manuell diese Clusters-Label angebe und sage, okay, diese Cluster, diese 100 Datensätze gehe ich durch und ich suche dann, ob die Cluster richtig alles gelabelt hat. Stimme ich zu oder nein? Und dann erstelle ich am Ende eine Labeled-Dataset, die aus dem Clustering-Modell A kommt. So eben für die Clustering-B, Clustering-C Modelle. Dann habe ich drei Labeled-Datasets. Drei verschiedene Versions von Datasets. Die nutze ich für die Klassifizierung vom Modell, für die Training von einem Klassifizierung Modells. Und dann am Ende, die Test-Errors vom Klassifizierungsmodell können sagen, welches Clustering Modell das erfolgreichste war. Weil die Erfolgrate von einer Klassifizierung wird stark von dem Struktur von Dataset beeinflusst. Und wenn die Dataset selbst nicht so richtig gelabelt ist dann Klassifizierung Modell kann nicht so mehr machen

A3

Verstehe. Das heißt, wenn ich das richtig verstehe, dann testest du selber deine Benchmarks und vergleichst du so wahrscheinlich auch verschiedene Clustering.

I9

Ja. Genau.

A3

Okay, interessant. Und das machst du, was würdest du dir wünschen, damit du das nicht mehr machen müsstest? Was würde dir helfen dabei?

I9

Ein Labeled Dataset von Anfang an. Das würde ich sagen. Aber ja, wenn ich dann wüsste, welches Clustering Modell, welches Klassifizierungsmodell am besten für so eine Anforderung die beste ist, dann würde das mir sehr helfen. Damit ich diese Analysephase ein bisschen reduzieren kann.

A3

verstehe ok dann würde ich sagen das ist eigentlich genau der punkt an dem unsere Studie ansetzt ich fange jetzt mal an dir was zu zeigen was wir uns denn so erarbeitet haben so ich hoffe man sieht meinen Browser ich habe schon gelernt hier unten kann ich das groß machen und zwar haben wir hier ein KI Label es hat nichts mit dem labeling von daten zu tun da gab es schon in der Vergangenheit Verwirrung aber es ist Label für ein KI Modell und ohne dass ich dich jetzt noch weiter beeinflusse würde ich gerne fragen was ist so dein spontaner Eindruck von diesem KI Label das du hier siehst?

I9

Also ich gehe davon aus, dass sie wollen für bestimmte Anforderungen die beste Modelle irgendwie herausfinden, die anpassenden Modelle.

A2

Ja, es ist vielleicht auch gar nicht, überleg mal, es ist nicht so wichtig, was wir jetzt uns damit gedacht haben, sondern wir haben jetzt wirklich Interesse daran, wenn du sowas siehst, was siehst du da? Was kommt dir jetzt in den Kopf? Was fällt dir auf? Was findest du vielleicht unklar? Ganz einfach mal ganz wild drauflos erzählen.

I9

natürlich es ist meine Frage sie sagen KI Labeling sie labeln nicht data sondern die Modelle selbst und es gibt verschiedene Models aber dann die Frage ist wozu? Die frage ist warum wollen sie einem Label [Schlechte Verbindung] ermittelt diesem Label

A3

Jetzt habe ich gerade den letzten Teil nicht verstanden. Wer ermittelt dieses Label?

I9

ja genau also ich wollte ich möchte wissen was ermittelte so ein label über das modell

A3

Okay, ja klar. XXXXXXXXX, meinst du, es wäre schon okay, wenn ich jetzt schon in die Erklärung...

A2

Absolut, ja. Dann ist das auch eine gute Antwort

A3

Na klar, alles klar. Dann erzähle ich gerne mehr darüber. Also die Idee dahinter ist natürlich erstmal diese Entscheidungsfindung zwischen verschiedenen Modellen, so wie du das eben selber machst, also das Benchmarking irgendwie schön darzustellen und somit eben die Entscheidungsfindung zu vereinfachen. In diesem Fall ist es so, dieses KI-Label entstammt einem Paper von unserem Kollegen XXXXXXX. Der hat sich verschiedene Bild Klassifikatoren, du siehst ja hier das ImageNet, angeschaut. In dem Fall ist es natürlich die Modellkarte für das MobileNetV3. Könnte dir auch was sagen. Diese KI-Modelle wurden miteinander verglichen bezüglich verschiedener Benchmarks. Ein paar siehst du hier unten. Also wir haben natürlich den durchschnittlichen Stromverbrauch pro Inferenz, also pro einmal Daten durch das neuronale Netzwerk schicken. Dann haben wir die Robustness, also wenn Attacken eben auf das, also Pixel-Attacken auf das Bild gefahren werden, wie robust ändert sich die Entscheidung des KI-Modelles in diesem Fall oder bis zu welchem Grad kann man gehen, damit die Entscheidung eines KI-Modells immer noch die gleiche bleibt, trotz Attacken. Dann haben wir natürlich die Top-1-Accuracy, also einfach das beste Ergebnis des KI-Modelles, in wie vielen Fällen dieses dem wahren Ergebnis der Klassifizierung eben entspricht. Und die Running-Time-Per-Inference haben wir hier auch noch angegeben. Es gibt aber, wie gesagt, noch weitere Benchmarks, die herangezogen werden, aus denen sich in Summe, in Gänze, eine große Klassifikation eben berechnet. Ja.

I9

Also Sie erstellen hier einen Vergleich.

A3

Ja, in der Tat. Ich würde es einfach mal machen, ich würde jetzt weitergehen und nochmal, damit es vielleicht klar wird, noch ein zweites KI-Label zeigen. Darum geht es natürlich. Kannst du mir vielleicht jetzt sagen, was du davon hältst und was dein Eindruck ist? Von diesen KI-Labels.

I9

Das ist natürlich eine sehr gute Idee. Was war ein Eindruck? Das ist mir ganz neu, was Neues und das finde ich super. Aber es kommt darauf an, welche Daten, also auf welchen Datasets haben Sie das schon mal probiert?

A3

Das wurde jetzt auf ImageNet einfach getestet. Also ImageNet ist ja dieser relativ große Bilddatensatz mit 1000 Klassen und es dient der Bildklassifizierung. Also letztendlich, ob eine Katze drauf ist oder ein Hund, 1000 verschiedene Klassen. Und das ist die Grundlage für diese Tests gewesen.

I9

Achso, so diese A, B, C, D Qualifikationen-Stufe, das ist ja. Das ist super. Und dann da bleibt die… Welche Gewichte zum Beispiel hat diese zum Beispiel Accuracy, der PowerDraw, um zu entscheiden, zu welchen Kategorien gehört EfficientNetB4. Zum Beispiel EfficientNet B4 gehört zu der Kategorie C, aber die Accuracy ist höher als die MobileNet.

A3

ja das ist vielleicht im ersten Moment irritierend aber es liegt daran dass ja noch weitere Accuracy… Sorry weitere Metriken dazu gezählt werden und die werden wenn ich XXXXXXX richtig verstanden habe gleichmäßig hinzugezogen also es sind insgesamt zehn Metriken also noch sechs weitere und jede Metrik geht zu einem Zehntel rein. Okay gut was gefällt dir an diesen Labeln vielleicht besonders was findest du bist gerade besonders hilfreich vor allem wenn du auch auf deine arbeitspraxis siehst also wenn du zwischen KI Modellen in deinem Job auswählen musst

I9

Also die Kriterien wie zum Beispiel Accuracy und Power, also diese Powerdraw per Inference ist mir wichtig wie zum Beispiel, weil bei meinem Unternehmen, wir haben internen Service, wir nutzen keine Cloud oder Cloud Services, wir haben unsere Service. Und deswegen einfach erstmal die Performance ist mir wichtig, ein Modell auszusuchen und zweitens, also erstens ist Accuracy, zweitens ist die Performance von diesem Modell, wie viel Strom braucht Server, um dieses Modell durchzugehen.

A3

und könntest du in dem Fall die beiden Labels also die beiden Modelle in dem fall werden es EfficientNetB4 und mobileNetv3 miteinander vergleichen und da auch eine Entscheidung treffen für deine Firma zum Beispiel? Würde dir das helfen?

I9

Nicht so dann ich muss in diesem Fall ich würde zum Beispiel sagen EfficientNetB4 würde mir die beste sein weil die Accuracy also das Gewicht von Accuracy von meiner Entscheidung ist mehr als Gewicht für den Power Draw und dann würde ich sagen also die linke Modell ist mir besser

A3

Ah, interessant. Okay, das linke Modell. Und dann, was würde für dich quasi dieser große Score hier für einen Einfluss haben auf diese Entscheidung?

I9

Das würde mir am Anfang nicht so wichtig sein. Also natürlich, es würde einen großen Eindruck an mir haben, aber wenn ich dann detailliert gehe und meine Anforderungen weiter analysiere, ich würde dann herausfinden, dass mir die Accuracy wichtiger ist, weil ich arbeite mit den Kunden. Und die Kunden, wenn ein Kunde falsch positiv klassifiziert ist als Betrüger, aber er ist nicht ein Betrüger, die Schaden natürlich ist nicht so gut zu ertragen. Deswegen, ich würde sagen, diese Kategorien A, B, C, D, E, würde bei meiner Entscheidung am Anfang nicht so wichtig sein, sondern die Accuracy ist wichtig, weil ich eine Kriterie hier habe. Und ich suche nach dem präziesestem Modell hier.

A3

Verstehe würdest du vor dem Hintergrund das Label irgendwie anders gestalten hast du eine Idee, wie das anders aussehen könnte damit es für dich noch hilfreicher wäre?

I9

dann ich würde erstmal Fragen wie entscheiden sie sich vor die Kategorie a und Kategorie c und bei einem Modell?

A3

also erstmal die Frage also das ist ein compound also einfach ein gewichtetes mittel letztendlich aus den zehn Metriken und natürlich ist es immer im Vergleich also es gibt immer Spektren…

I9

Woher kommt diese gewichte für diese Metriken?

A3

In dem Fall ist ja eine Gleichgewichtung jedes jede Metrik ist gleich gewichtet

I9

In diesem Fall, bei einigen Anforderungen, ich würde sagen, nicht alle Metriken sind gleich gültig. So wie bei meinen Anforderungen, [zum Beispiel, Accuracy-Statik. Und in diesem Fall, diese Kategorie, also die Labels, helfen, um eine Entscheidung zu treffen. Sondern nur eine Metric, nur eine bestimmte Metric.] (schlechte Verbindung)

A3

ok und wie würdest du das Label noch mal würdest du es konkret ändern dass diese Bedingung erfüllt ist welche Informationen sollten noch…

I9

Ich würde mehr Gewicht auf Accuracy legen und dann würde ich EfficientNetB4. Für mich würde A sein und MobileNetV3, würde zum Beispiel B oder C sein, soweit ich Ihre Frage verstanden habe.

A2

Ich glaube, das ist ganz richtig verstanden, aber sollte man jetzt visuell etwas ändern? Also sollte man die Informationen vielleicht anders präsentieren? Weil was du vorschlägst, das ist ja so wie so ein innerer Prozess. Das Label könnte gleich bleiben, es würde sich aber die Kategorie verändern. Wäre zum Beispiel für dich wichtiger, dass die Symbole unten größer sind und gar nicht so stark diese A, B, C, D, E…

I9

[Schlechte Verbindung] dass die Gewichte erwähnt werden.

A3

Ah, okay.

I9

Es ist aufwendig, was ich mir vorstelle, aber es wäre besser, wenn man wenigstens eine Permutation über die verschiedenen Gewichte hatte. Wenn wir zum Beispiel sagen wir haben zum Beispiel drei Stufen von Gewichten, Gewichte 1, 2 und 3. Und jede Parameter, Funktion von dieser ohne Metriken, zum Beispiel einmal sagen wir mal, Accuracy hat Gewicht 3, alle anderen Gewicht 1. Um so eine Permutation durchzugehen und verschiedene Labels für diese Permutation... Weil diese Entscheidung, soweit ich verstanden habe, diese A, B, C, D, E Kategorien stammt aus gleich gültigen Metriken. Also Zusammenkommen von den gleichgültigen Metriken. Aber wenn wir mehr Gewicht auf eine von diesen Metriken legen, dann würde diese Kategorisierung enden. Ist das richtig?

A3

Ja, das ist richtig. Also du hast es auf jeden Fall richtig verstanden, wie der Score entsteht. Würdest du, wie würdest du das Umsetzen mit den verschiedenen Permutationen? Jetzt visuell, wie soll es präsentiert werden? Würdest du dir zum Beispiel drei verschiedene, also pro Permutation ein großer Score oder wie stellst du dir das vor?

I9

Ja, es ist ein [schlechte Verbindung] im Design und der Vorstellung, aber wie würde ich das [Schlechte Verbindung] schwer auf, aber ich finde... Nächsten sagen die gleichgültigen Metriken zusammen einmal eine Metrik als die wichtigste, alle anderen gleichgültig. Und dann viermal ABCD, ABCD, ABCD. Je nachdem, die beste. Also je nachdem, die wichtigste Metrik und eine da oben, die alle gleich gültige Metriken.

A3

Okay, danke. Das kann ich mir jetzt besser vorstellen. Wie würde sich das denn verhalten zu diesen Metriken hier? Weil die haben ja quasi schon, die sagen ja schon, das ist die eine wichtige Metrik. Also daran kannst du es ja auch schon ablesen.

I9

Also die Farben. Das sagt schon, ja. Und deswegen habe ich am Anfang gesagt, ich würde für die linke Model entschieden, weil Accuracy ist mir die wichtigste. Ich muss dann ja sagen, ich habe nichts zu sagen. Die Labeling-Verfahren, die sie schon ausgesucht haben, das ist ja die beste, dass ich mir vorstellen kann.

A3

Okay, alles klar. Dann trotzdem danke, dass du versucht hast, nochmal das kreative Gehirn anzuschmeißen. Ich hätte noch die Frage, jetzt wo du das Konzept verstanden hast, dass wir Modelle labeln und nicht Daten labeln, kennst du noch andere Kommunikationsformen, die genau das machen, also eben Modelle darstellen?

I9

Sorry, das habe ich nicht so völlig verstanden. Kommunikationsmodell?

A3

Ja, Kommunikationsformen, also Arten wie KI-Modelle, dir schon präsentiert worden sind.

A2

Also du hast ja zum Beispiel gesagt, du nutzt TensorFlow, also du gehst in irgendwelche Libraries. Wie findest du Informationen über die Modelle, mit denen du arbeitest? Ist da irgendwie eine Documentation oder wie kannst du das greifbar machen?

I9

Also, [Schlechte Verbindung] lese ich diese Fachtexte [Schlechte Verbindung] und dort lese ich die Einträge. Je mehr Informationen wir kriegen, und halt Clustering, AP, Modelle sind mir geeignetste. Und dann gehe ich dann in GitHub und lese alle andere Beiträge und Quellcodes von der Anforderung. Aber sie meinen, so habe ich es verstanden, sie meinen, wie meine Arbeitsverfahren

A3

Ja, also wie du die Informationen reinholst, über welche Arten. Ganz kurz ein Einschub. Könntest du vielleicht dein Bild ausmachen, weil es war jetzt gerade wieder sehr hakelig, leider. Ja, das gerade war gut. Können wir es auch sehen? Also ich sehe dich, genau. Du kannst aber auch gerne sonst die Kamera ausmachen, weil wenn es dem Ton gut tut, dann tut es auch unserem Interview, glaube ich, und gerade der Transkription besser. Ja, danke schön. Und jetzt hört es sich auch schon besser an, habe ich das Gefühl. Okay, du hast es richtig verstanden. Also die Arten, wie eben du dir Informationen über die KI-Modelle holst. Das hat XXXXX ja auch schon richtig angesprochen. Da haben wir eine kurze Zusammenfassung für dich. Wir haben gefunden, wie man sich Informationen holen kann über verschiedene KI-Modelle, gerade wenn man in solchen Entscheidungssituationen steht. Die gehen wir jetzt mal ganz kurz durch. Also wir haben natürlich den Klassiker, ist der auch nicht unbekannt. Wissenschaftliche Publikationen, zum Beispiel beim MobileNetV3, was wir ja gerade schon in der Labelkarte hatten. Das ist natürlich aus einer wissenschaftlichen Publikation, aus einem Paper, in dem Fall, von einem Google-Team entstanden. Und da kann man sich natürlich wirklich das Modell wirklich im größten Detailgrad sich anlesen und man hat auch Benchmarks dazu. Entschuldigung. Dann sehen wir hier rechts die Model-Card. Das ist von Google, müsste das sein. Ich glaube, das ist von TensorFlow in dem Fall. Auch hier wieder für das MobileNetV3, Small, wo gewisse Informationen zu zum Beispiel Input-Daten, und auch der Anzahl von Parametern im Modell zu sehen sind. Wie groß ist das Modell überhaupt? Und natürlich auch Benchmarks. Also hier sehen wir Inference-Zeiten. Und natürlich die Bedingungen für die Benchmarks sind hier aufgezeichnet. Dann haben wir Papers with Code. Ich weiß nicht, ob dir das schon begegnet ist. Eine sehr schöne Seite, bei der man sich eben ansehen kann, welche Paper denn ein gewisses KI-Modell zum Beispiel zitiert haben. In dem Fall haben wir hier wieder MobileNetV3 und wenn man jetzt hier runter scrollen würde auf dieser Webseite würde man alle wissenschaftlichen Publikationen einsehen können die MobileNetV3 in irgendeiner Weise zitiert haben. Dann natürlich, das hast du glaube ich gerade auch schon gesagt, dass du dir das durchliest, also so ein Blog-Artikel, Towards Data Science oder auch Medium.com, sind sehr beliebt. Und da geht es dann auch eben, da gibt es viele Informationen zu den einzelnen Modellen. Und zu guter Letzt natürlich noch die Libraries, klar, also aus der Python-Dokumentation oder aus TensorFlow, PyTorch, die natürlich die technischen Spezifika vor allem enthalten, also welche Argumente kann ich dem Modell geben und auch hier wieder Benchmarks zu sehen, ganz unten. Und zu aller guter Letzt haben wir noch die Factsheets. Das ist von IBM. Das gibt es also auch nur für IBM produzierte KI-Modelle. Und die enthalten verschiedene Kategorisierungen, wie eben die KI-Modelle beschrieben werden können. Also einmal zum Beispiel grobe generelle Informationen, wie groß ist das Modell, aber auch welche Inputs und Outputs gibt dieses Modell. Dann aber auch Sachen zu wirklich Biases, also wie fair ist das Modell für gewisse Klassifikationen und auch Robustheitstests. Das ist alles sehr ausführlich. Man muss aber dazu sagen, das ist nur für IBM-Modelle leider. Das war's. Das sind die anderen Arten von Kommunikation, die wir gefunden haben. Zunächst einmal, welche Vor- und Nachteile siehst du? Zwischen unserem, dem dir gerade vorgestellten Modell und den verschiedenen Formen der Modellkommunikation.

I9

so entschuldigt deine Frage habe ich nicht verstanden kannst das wiederholen?

A3

Die Frage ist, welche Vor- und Nachteile siehst du zwischen den verschiedenen Darreichungsformen insbesondere zwischen dem Label…

I9

natürlich die Zeitaufwand und Umfang also die vergleichen wie zum Beispiel [Schlechte Verbindung] …

ImageNet selbst oder von mobileNet selbst die haben verschiedene Datasets benutzt um Benchmarking durchzugehen aber in diesen Labels die wir haben ganz am rechts es geht um nur eine jetzt. Zusammengefasst Zeitaufwand umfangreiche vergleichstellung benchmarking also die die vielfältigen benchmarking datasets sind die zwei Kriterien die ich [Schlechte Verbindung]

A3

Okay. Okay, alles klar. Vielen Dank. Dann, okay, noch irgendwelche weiteren Nachteile, die dir einfallen? Also Nachteile wäre quasi, dass wir, oder dass im AI-Label hier rechts, mehrere Benchmarking-Datensätze benutzt werden sollten. Habe ich das richtig verstanden?

A2

XXXX lass uns auch mal die Kamera ausmachen, dann ist nicht so viel incoming.

I9

Können Sie mich besser hören jetzt?

A2

Also ich, jetzt wird es hoffentlich ein bisschen noch mal besser.

A3

Ja.

A2

Kannst du noch gleich nochmal den letzten Punkt...

A3

Ich muss mal ganz kurz das Teilen beenden.

I9

Hören Sie mich jetzt gut?

A3

Ja. Ja, jetzt gerade hört sich das sehr gut an.

A2

Gar kein Problem. Wir wollen nur, dass wir Ihre Zeit jetzt auch benutzen, damit wir Ihre Antworten nachher gut auswerten können. Es ist ja wichtig, was Sie sagen. Daher können Sie nochmal einfach vielleicht im Gesamtpaket wiederholen. Also Zeitaufwand, Aufwand habe ich verstanden. Der letzte Punkt, der ist mir nicht ganz klar, der Nachteil.

I9

Nicht Nachteile, also ich würde noch mal wiederholen. Erstens, die Zeit auf alle anderen Vorgehensweise, wie zum Beispiel diese durch die Publikationen durchzulesen oder die anderen Internetbeiträge und sowas, alles durchzugehen, das ist zeitaufwendig. Genau, ja. Das ist der Nachteil von allen anderen Vorgehensweisen, aber der große Vorteil, den Ihr Label hier hat, ist, dass es mir hilft, sofort zu einer Entscheidung zu landen. Also es reduziert viel diese Zeitaufwand. Die anderen Vorteile, die Ihr Label hier hat, ich würde sagen, ja, die umfangreichen Vergleiche, sorry, das würde ich eigentlich wegpacken aus meiner Entscheidung, aber der Nachteil, einen Vorteil habe ich erwähnt, der Nachteil bei Ihrem Label ist diese Datasets, also die Benchmark-Datasets. Sie haben ihren Vergleich basierend auf einem Dataset durchgeführt. Es wäre besser, wenn mehrere Benchmark-Datasets hier berücksichtigt werden.

A3

Okay, noch eine Frage, vielleicht eine letzte Frage zu diesem Vergleich. Wie siehst du das mit der Dokumentation, also die MobileNetV3 Small Dokumentation. Siehst du hier noch einen entscheidenden Vor bzw. Nachteil zwischen Dokumentationen, also wirklich das, was du beim Programmieren benutzt und dem Label?

I9

als Programmierer ich würde mir wünschen das Modell schon ganz flexibel und robust auf dem GitHub zur Verfügung steht. Es wäre besser dass so eine Information dieses Label mir ermitteln würde ob diese Modelle vorhanden ist wissen wir also ich kann sagen zum Beispiel es gibt viele Modelle die Publikationen für den ganze offen da draußen sind aber die code diese Quellcode nicht im GitHub vorhanden sind. Kann man zum Beispiel die Quellcode vor einem Modell finden das war zum Beispiel mir eine große Herausforderung als ich mal an meiner Abschlussarbeit gearbeitet habe

A3

Ja, das ist eine gewisse Intransparenz, das ist auch ein gutes Stichwort. Die Frage wäre, wer sollte deiner Meinung nach dieses Label ausstellen? Also wer sollte zertifizieren?

I9

ich glaube ganz demokratisch, die Benutzer

A3

Interessant. Kannst du da nochmal drauf eingehen, wie genau?

I9

Es könnte sein, dass eine zentrale Website oder Stelle wäre. Also ich kann es mir chwer vorstellen, aber ich kann mir, wie zum Beispiel Towards Data Science. In Towards Data Science zum Beispiel, die Benutzer können einfach sagen, ja, diese Labels haben mir viel geholfen oder diese Labels.

A2

Naja stell dir vor… Vielleicht hilft es so: Also jetzt haben wir das ja bzw. XXXXXXX von der XX XXXXXXXX das erstellt, dann könnte man sagen die XX XXXXXXXX hat das gemacht. Da ist jetzt die Frage ist das gut oder ist das schlecht? Also wer sollte das jetzt zur Verfügung stellen? Also in diesem Fall hat das Ja XXXXXXX entwickelt von der XX XXXXXXXX, da könnte man sagen hier gibt es jetzt diese Label von der XX XXXXXXXX, aber es gibt ja auch andere Möglichkeiten. Also ein großes Unternehmen, was ja zertifiziert ist, zum Beispiel TÜV, genau, es könnte aber ja natürlich auch irgendwie ein kommerzielles Unternehmen sein.

I9

dazu also es wäre auch jetzt… also ich verstehe die Frage aber ich verstehe die ganz… die Konzepte hinter dieser Frage nicht

I9

Geht das um Marketing?

A2

Nee, gar nicht. Also stell dir vor, du hast jetzt deine Aufgabe, du willst ein Clustermodell finden und du hast jetzt dieses Label und willst anhand dessen eine Entscheidung treffen. Jetzt ist ja die Frage, vertraust du dem oder nicht? Und was würde dir helfen, dass das irgendwie vertrauenswürdig ist? Also wenn das irgendwie von der Uni kommt oder wenn das jetzt die Firma AI Label GmbH erstellt hat. Wäre das für dich wichtig in deiner Entscheidung? Oder würdest du sagen, ja, das ist jetzt halt so ein Label, das finde ich schon gut?

I9

Achso, kann ich aus Industrie nehmen. Und wenn zum Beispiel so ein Institut wie zum Beispiel TÜV mir sagt, diese Labeling sind schon geprüft, das würde mir mehr vertrautheit vermitteln. Weil aus der Universität manchmal die Fragen, also die Probleme, die in Universitäten zu lösen sind, sind anders als in der Industrie.

A2

Absolut, ja.

I9

Und deswegen würde ich auf eine andere externe Organisation wie zum Beispiel TÜV mehr anvertrauen.

A2

Ja, könnte aber ja zum Beispiel auch irgendwie ein politischer Akteur sein, also könnte ja jetzt auch die EU machen, wie fändest du sowas?

I9

EU wäre besser. Aber ja, so ein Labeling zu erstellen, aber die Vertrautheit würde erstmal nicht eine große Rolle spielen. Weil man als ein Maschineningenieur, man weiß zu einem gewissen Punkt, dass dieses Modell zu diesen Problemstellungen schon die beste Leistung beigetragen hat. Und natürlich würde es mir nicht so wichtig sein, wer dieses Label erstellt halt.

A3

Okay. Das heißt, weil du viel weißt, wäre dir das nicht so wichtig, weil du von dir aus recherchierst.

I9

So kann man jetzt das interpretieren, genau.

A3

Ja, okay. Auch spannend, ja. Okay, dann hätte ich noch abschließend die Frage, wir haben jetzt viel darüber geredet, wie das denn jetzt wirklich gestaltet werden könnte. Wie hilfreich fändest du diese Zertifizierung, vor allem im Hinblick auf deinen Arbeitsalltag, wenn es sowas geben würde?

I9

Es würde mir sehr hilfreich sein, wenn schon mal so ein Labeling schon mal über die detaillierten Anforderungen geht, nicht nur die benchmarking-Datas. Was kann ich sagen? Das wäre natürlich sehr hilfreich. Das kann mir dabei helfen, viel Recherchieraufwand zu reduzieren. Aber die Frage ist, ob diese Labels aktualisiert sind. Es kann sein, dass Labels gehören zum letzten Jahr und mittlerweile gibt es schon zehn oder zich neue Modelle. Und dieses Label ist ein bisschen alt geworden und deswegen kann ich mich nicht mehr darauf verlassen.

A3

Ja, dazu haben wir ja hier zum Beispiel auch ein Feld. Würde das auch Vertrauen quasi aufbauen? Oder verbessern?

I9

Ja.

A3

Okay. Dann noch die allerletzte Frage. Müsste das Label noch irgendwie angepasst werden an deinen Wissenshintergrund? Wünschst du dir vielleicht mehr, weil du Machine Learning Professional bist?

I9

Entschuldigung, kannst du die Frage nochmal …

A3

Würdest du dir mehr Informationen wünschen auf dem Label?

I9

mehr informationen ich muss jetzt ein bisschen nachdenken was war ein weiter wir haben jetzt

A2

Also wir haben ja zum Beispiel, siehst du jetzt hier nur so vier Benchmarks, aber es könnten ja zum Beispiel auch alle zehn drauf sein.

I9

Ja, verstehe. Also die Aktualisierungsdatum, Modell, Benchmarkendatasets, Kategorie. Also diese Framework TensorFlow, das könnte auch sein. Man kann auch zum Beispiel andere Frameworks hier nutzen, wie zum Beispiel PyTorch. Viele Unternehmen arbeiten mit bestimmten Text-Stacks. Zum Beispiel ein Unternehmen arbeitet nicht mehr mit TensorFlow, sondern PyTorch. Es könnte auch besser sein, wenn eine Information über die Frameworks, mehr Informationen über die Frameworks auf dem Label steht. Dann nehmen zum Beispiel TensorFlow auch PyTorch.

A3

okay vielen dank ich hätte keine weiteren fragen mehr XXXXX hättest du noch weitere Fragen?

A2

Auch nicht. Ich fand es sehr, total spannend. Und ich habe, glaube ich, nochmal ganz, ganz neue Einblicke bekommen. Also ganz herzlichen Dank, XXX. Ich würde meine Kamera nochmal kurz anmachen und die Aufnahme stoppen.