

Persoonlijk Logboek - Master Applied AI

Blok 1: Beeldverwerking - Competentie Evaluaties

Projectreferenties

Project	Beschrijving	Link
Churn Prediction Pipeline	Afstudeerstage - Geautomatiseerde pipeline met 6 ML modellen	Research Article (Preprint)
TNO Speaker Recognition	CNN-gebaseerd speaker recognition model	Portfolio Details

Bewijs van Bijdrage

Bewijs ID	Beschrijving	Type	Locatie/Link	Gerelateerde Competenties
BEW-001	Gebruikersinterview transcript (AH medewerker)	Document	Document kan geleverd worden naar aanvraag i.v.m. gevoelige gegevens	A1, A2, A4
BEW-002	Happy Flow diagram	Mermaid	happy-flow.mmd	A3, C1
BEW-003	Error Flow diagram	Mermaid	error-flow.mmd	A3, C1
BEW-004	Preventief dataset onderzoek	Document	preventief-onderzoek-datasets.md	B1
BEW-005	Business Critical Zones scatterplot	Afbeelding	model-performance.jpg	B2, C2
BEW-006	Eigen paper prototype (Figma)	Online	https://green-frame-34768810.figma.site/	A3, C3
BEW-007	Technische Review Feedback (groep)	Document	feedback-technisch-rapport.md	B1, B2, B3, C2
BEW-008	M&T Opdracht 1a Feedback	Document	feedback-mt-opdracht1a.md	A2, B1, B2, C2
BEW-009	PvA Artikel Feedback	Document	feedback-pva.md	D2, D3

A. CONTEXTUALISEREN EN ONTWERPEN

A1: Onderzoek en herformulering van vraagstukken

S - Situatie: Tijdens mijn afstudeerstage heb ik deze competentie ontwikkeld door vanuit de vraag "wat kunnen we met onze huidige data doen" een concrete oplossing voor te stellen: een geautomatiseerde churn prediction pipeline. Dit proces vereiste het herformuleren van een vaag vraagstuk naar een technisch haalbare oplossing, waarbij ik de belangen van verschillende stakeholders (CTO, CEO, eindgebruikers) heb afgewogen. De stage resulteerde in een pipeline met 6 verschillende machine learning modellen, 3 temporal validation strategies, en 3 resampling techniques. Door het proces heen heb ik regelmatig meetings gehad met management om bevindingen en experimenten te koppelen aan business concerns en praktische relevantie te behouden. Mijn sterke punten liggen in oplossingsgericht, praktisch, en pragmatisch denken. Mijn grootste uitdaging binnen deze competentie is het hele project vanaf het begin duidelijker in kaart brengen, zodat fouten en onduidelijkheden eerder in het proces voorkomen kunnen worden.

T - Taak: Voor het leeftijdscontrole project ga ik helpen om het project duidelijk in kaart te brengen via bijvoorbeeld empathy maps, value proposition, gebruikersinterviews en haalbaarheidsonderzoek.

A - Actie: Ik heb een gebruikersinterview afgelegd met een medewerker van Albert Heijn (hoofdkantoor) (rol: medewerker bedrijf). Het interview heb ik voorbereid met gestructureerde vragen om dieper in te kunnen gaan op interessante punten. Daarnaast heb ik via dit contactpersoon (een goede vriend) ook buiten het formele interview om verduidelijkende vragen kunnen stellen. (Bewijs: BEW-001)

R - Reflectie: Het interview verliep soepel doordat ik de geïnterviewde al goed kende, wat een open gesprek makkelijker en meer toegankelijk maakt. Ik merkte dat mijn ervaring met stakeholder interviews tijdens mijn afstudeerstage me hielp om doorvragende vragen te stellen. Een verbeterpunt is dat ik meer gebruikers had kunnen interviewen om een breder perspectief te krijgen. Ik heb nu vooral het perspectief van personeel, maar minder gericht onderzoek van de klanten zelf (alleen een korte enquête).

T - Transfer: Voor de rest van het blok ga ik de opgedane inzichten gebruiken bij het verfijnen van requirements en het testen van ons prototype. Voor toekomstige projecten neem ik mee dat het waardevol is om vroeg in het proces contact te leggen met stakeholders, en dat het hebben van een vertrouwenspersoon binnen de doelgroep het onderzoek vergemakkelijkt.

A2: Requirements opstellen

S - Situatie: Voor het churn prediction project heb ik juridische, organisatorische en technische vereisten gedocumenteerd in mijn research article. Tijdens mijn bachelor heb ik uitgebreide ervaring opgedaan met functionele en technische eisen via MoSCoW-analyses in verschillende projecten. Daarnaast heb ik een ethisch en juridisch verslag geschreven over Deep Fake & AI technologieën, wat mij een basis heeft gegeven in de ethische en juridische aspecten van AI. Mijn sterke punten liggen in het systematisch documenteren van technische en organisatorische requirements. Het gebied waar ik minder ervaring heb is de ethische kant van requirements opstelling.

T - Taak: Requirements opstellen voor het leeftijdscontrole project, inclusief juridische (AVG, AI Act), ethische (bias), organisatorische (workflow), functionele en technische vereisten.

A - Actie: Ik heb gezamenlijk met mijn team requirements opgesteld door AVG-wetgeving en AI Act te analyseren, ethische aspecten te bespreken tijdens filosofie & ethiek lessen, functionele requirements af te leiden uit de gebruikersinterview, en technische requirements te bepalen op basis van literatuuronderzoek. (Bewijs: BEW-001, BEW-008)

R - Resultaat: We hebben een uitgebreide requirements lijst opgesteld die in ons technisch rapport staat. De technische review gaf aan dat de requirements "prima" zijn en dat we over veel aspecten hebben nagedacht (beoordeling: "Op niveau"). Echter bleek er een misverstand tussen 8 MUST haves in de bijlage versus 4 in hoofdstuk 4. (Bewijs: BEW-007) Verder, in mijn individuele M&T feedback bleek dat niet alle requirements voldoende uit de opdracht of data-analyse voortvloeien. (Bewijs: BEW-008)

R - Reflectie: Ik ben tevreden met de breedte van onze requirements: we hebben juridische, ethische, organisatorische, functionele én technische aspecten opgezet. Ik hoop in de toekomst meer tijd te spenderen aan de ethische en functionele eisen, want ikzelf heb voornamelijk de focus op de overige requirements gelegd.

T - Transfer: Voor het eindrapport moet elke requirement expliciet gekoppeld worden aan een bron (interview, wetgeving, literatuur, data-analyse). Voor toekomstige projecten neem ik mee dat requirements niet alleen moeten worden opgesteld, maar ook moeten worden geverifieerd en gevalideerd met stakeholders en bronnen. Verder zal een hogere mate van focus voor ethische en functionele eisen nodig zijn voor mij.

A3: AI-specifieke ontwerprichtlijnen

S - Situatie: Bij mijn afstudeerproject heb ik het CRISP-DM framework toegepast om de hele lifecycle van het project te bewaken en te structureren. Aspecten zoals veiligheid, privacy en error handling heb ik in verschillende projecten aangepakt als vanzelfsprekende onderdelen van de ontwikkeling. Echter, specifieke richtlijnen zoals vertrouwen, feedback mechanismen en controlemechanismen heb ik nog niet expliciet geïmplementeerd of gedocumenteerd.

T - Taak: AI-specifieke ontwerprichtlijnen toepassen voor vertrouwen, veiligheid, privacy, elegante foutafhandeling, feedback en controlemechanismen in het leeftijdscontrole project.

A - Actie: Ik heb happy en error flows ontworpen die de basis vormden voor onze prototypes. De error flow specificeert wat er gebeurt bij verkeerde leeftijdsschattingen en edge-cases. Deze flows zijn gemaakt met Mermaid en geïmplementeerd in ons prototype. (Bewijs: BEW-002, BEW-003, BEW-006)

R - Resultaat: De flows vormden een basis voor hoe het systeem moet reageren op verschillende scenario's. Ze evolueerden van paper prototypes naar ons huidige werkende prototype. De error flow heeft specifiek geholpen om na te denken over elegante foutafhandeling en transparantie richting gebruikers.

R - Reflectie: Het ontwerpen van flows dwong me om expliciet na te denken over wat er misgaat in plaats van alleen de happy path. Dit is een mindset shift die ik waardevol vind voor AI-systemen. Mijn achtergrond in systematisch denken (van mijn bachelor en afstudeerstage) hielp hierbij. Een verbeterpunt is dat ik meer AI design patterns (zoals die van Google's People + AI Guidebook) expliciet had kunnen toepassen en benoemen in plaats van intuïtief te ontwerpen.

T - Transfer: Voor het eindrapport ga ik de ontwerpen documenteren met expliciete verwijzingen naar design patterns die we hebben toegepast. Bij toekomstige projecten zal ik eerder systematische design methodologies raadplegen in plaats van enkel op intuïtie te vertrouwen. De error flow benadering neem ik mee naar elk AI-project: begin altijd met "wat kan er misgaan?"

A4: Geschiktheid van AI onderbouwen

S - Situatie: Deze afweging heb ik gemaakt bij het voorstellen van mijn afstudeerproject, waar ik samen met de CEO en CTO onderzoek deed naar haalbaarheid en een MVP opzette. Het proces omvatte technische analyses en haalbaarheidsonderzoek, maar lag voornamelijk gefocust op technische aspecten. De maatschappelijke en ethische overwegingen bij het inzetten van AI waren niet expliciet gedocumenteerd of uitgebreid besproken. Dit is een competentie waar ik graag een meer holistische benadering wil ontwikkelen die technische, maatschappelijke en ethische aspecten gelijkwaardig behandelt.

T - Taak: De geschiktheid/haalbaarheid van AI voor leeftijdscontrole onderbouwen vanuit technische en ethische aspecten, en de gewenste mate van automatisering bepalen.

A - Actie: We hebben regressie versus classificatie als architectuurkeuze vergeleken. Bij regressie krijg je een continue leeftijdsschatting (bijv. 23.4 jaar), bij classificatie een discrete klasse (bijv. "<18", "18-25", ">25"). Na afweging hebben we gekozen voor regressie wegens hogere mate van inzichtelijkheid: we kunnen de exacte leeftijd die het model voorspeld inspecteren en zo zien in welke ranges het model meer of minder accuraat is.

R - Resultaat: De keuze voor regressie bleek achteraf goed omdat het ons de flexibiliteit gaf om business critical zones te definiëren (zie B2). We kunnen nu verschillende risico-niveaus visualiseren en het systeem kan worden aangepast zonder het model opnieuw te trainen op andere thresholds.

R - Reflectie: De keuze voor regressie is een keuze waar ik zelfverzekerd in sta: het was in deze fase van het project de meest logische keuze, met weinig nadelen

T - Transfer: Voor het eindrapport ga ik deze afweging explicieter maken en verschillende human-in-the-loop scenario's uitwerken. In toekomstige projecten neem ik mee dat architectuurkeuzes (zoals regressie vs classificatie) niet alleen technische gevolgen hebben maar ook fundamentele gevolgen voor de mate van automatisering en verantwoordelijkheidsverdeling. Later in het project kunnen we classificatie nog toevoegen als vergelijkingspunt.

B. ONTWIKKELEN EN MODELLEREN

B1: Dataset exploratie en preparatie

S - Situatie: Ik heb bij verschillende projecten datasets van scratch geanalyseerd, getransformeerd en ingezet voor machine learning modellen. Dit varieert van het preprocessen van volledig nieuwe datasets tot het fine-tunen van bestaande modellen op gespecialiseerde datasets, zoals bij het TNO Speaker Recognition project. Voor dat project ontwikkelden we een CNN-architectuur waarbij we een eigen dataset moesten prepareren en optimaliseren voor speaker recognition doeleinden. Mijn ervaring bestaat uit zowel traditionele machine learning datasets als meer complexe auditory en visual data preprocessing. Ik ben comfortabel met de technische aspecten van data preparation/preprocessing, maar wil graag meer diepgang ontwikkelen in o.a. de implicaties van data bias.

T - Taak: Geschikte datasets voor leeftijdsschatting verkennen en prepareren, rekening houdend met technische en ethische randvoorwaarden zoals bias en representativiteit.

A - Actie: Ik heb preventief onderzoek uitgevoerd naar beschikbare face age estimation datasets waarbij ik o.a. UTKFace, IMDB-WIKI, CACD, FairFace, en Lagenda heb vergeleken op criteria als grootte, leeftijdsverdeling, etniciteit-diversiteit, kwaliteit, en facial expression diversity. Dit onderzoek leidde tot de keuze voor UTKFace en IMDB-Cleaned als primaire datasets, met Lagenda als potentiële aanvulling indien meer data nodig is. Daarnaast heb ik data augmentation technieken verkend (greyscaling, rotation, shifting,

flipping) en de data loading code voor beide datasets geïmplementeerd. (Bewijs: BEW-004) Opmerking: De implementatie van data loading en preprocessing staat in onze private GitLab repo en kan niet extern als bewijs worden toegevoegd.

R - Resultaat: De technische review beoordeelde ons werk als "Boven verwachting" en merkte op: "De datacleaning is goed uitgevoerd en er is een mooie gecombineerde dataset ontstaan. Jullie zijn er van bewust welke 'fouten' onbalansen er nog in de gecombineerde dataset zitten." (Bewijs: BEW-007) We hebben een werkbare gecombineerde dataset van ~60K afbeeldingen met diverse leeftijden (0-100 jaar) en redelijke etniciteit-spreiding. We zijn ons bewust van resterende onbalansen: ondervertegenwoordiging van oudere leeftijden (70+) en Aziatische etnicity. Bij de opdracht van M&T was echter de feedback dat de data exploratie op de gehele dataset was gedaan zonder train/test splitsing, wat data leakage kan veroorzaken. (Bewijs: BEW-008)

R - Reflectie: Het preventieve literatuuronderzoek naar datasets hielp om een geïnformeerde keuze te maken voordat we begonnen met implementeren. Dit bespaarde tijd omdat we niet achteraf datasets hoefden te wisselen.

Echter, de M&T feedback over data leakage was helaas een fundamentele fout. Deze fout had makkelijker voorkomen worden met wat meer tijd voor de opdracht. Soms duik ik te snel in de praktijk zonder alle methodologische stappen zorgvuldig te doorlopen.

T - Transfer: Bij een volgende opdracht van M&T kan ik de volgende guidelines volgen:

- Split data (train/test, eventueel val)
- Dan pas EDA op training set
- Preprocessing fitted op training set
- Test set blijft "unseen" tot finale evaluatie

Voor het artikel ga ik ook explicieter ingaan op de ethische trade-offs van onze dataset keuzes en de resterende biases. Dit is relevant voor D2/D3.

B2: Architectuur en AI-technieken selecteren

S - Situatie: Deze competentie heb ik in verschillende contexten toegepast. Voor machine learning is mijn afstudeerproject een voorbeeld waar ik 6 verschillende algoritmes vergeleek en selecteerde. Voor deep learning ontwikkelden we een CNN-architectuur voor het TNO Speaker Recognition project. Daarnaast heb ik ervaring met computer vision via een movie poster classification model en natural language processing via een RAG-pipeline chatbot voor Inholland studiehandleidingen. Het gebied waar ik nog geen ervaring mee heb is kennisrepresentatie. Mijn sterke punt ligt in het praktisch vergelijken en implementeren van verschillende technieken.

T - Taak: Geschikte architectuur selecteren voor leeftijdsschatting gebaseerd op computer vision technieken, requirements, en beschikbare data.

A - Actie: Voor het technische rapport heb ik een VGG16-gebaseerd CNN model ontwikkeld en geëvalueerd. Als evaluatiemetriek heb ik MAE (Mean Absolute Error) als primaire metriek gebruikt omdat deze intuïtief interpreteerbaar is: hoeveel jaar zit het model er gemiddeld naast? Belangrijk waren verder de business critical zones: ik heb visualisaties gemaakt die praktische relevant inzicht en specifieke risico-percentages geven:

- % mensen jonger dan 18 die als 25+ worden voorspeld (hoog risico: illegale verkoop)

- % mensen 25+ die als <25 worden voorspeld (laag risico: gemaksfactor; kan klanttevredenheid verminderen)

(Bewijs: BEW-005, BEW-007)

R - Resultaat: De technische review beoordeelde dit als "Op niveau" en gaf aan: "Model beschrijving in het document is wat summier maar klopt wel, tijdens presentatie goed toegelicht." Er was feedback dat we hadden moeten beginnen met beschrijving van een basismodel (DummyRegressor) en dan de verbeteringen daarop beschrijven. (Bewijs: BEW-007) De business critical zones werden specifiek uitgelicht: "De Business-Critical Zone Analysis voegt veel waarde toe en maakt inzichtelijk hoe het model presteert voor de belangrijkste groep."

R - Reflectie: Een sterk punt was het koppelen van evaluatie aan business/praktische consequenties met de business critical zones. Dit idee kwam voort uit het besef dat een absolute error van 5 jaar heel verschillend is als het om een 15-jarige gaat (voorspeld als 20) versus een 40-jarige (voorspeld als 45). Echter, de feedback dat de model beschrijving "summier" was en we beter een baseline model hadden kunnen beschrijven laat zien dat correcte documentatie even belangrijk is. Ik ben vaak sterk in het bouwen maar minder sterk in het documenteren van de keuzes achter het bouwen.

T - Transfer: Voor het eindrapport ga ik in meer detail het proces van DummyRegressor tot uiteindelijke model beschrijven.

B3: Systematische modelontwikkeling

S - Situatie: Tijdens mijn afstudeerproject heb ik het CRISP-DM framework gebruikt als leidraad voor systematische en iteratieve modelontwikkeling. Dit proces bevatte regelmatige stakeholder feedback momenten en aanpassingen gebaseerd op business requirements en technische bevindingen. De structuur was cruciaal omdat het een significant project was met verschillende combinaties van modellen, validatiestrategieën en resampling technieken die allemaal getest moesten worden om de beste configuratie per bedrijf/uitgever te vinden. Ik ben sterk in het volgen van gestructureerde methodologieën en het documenteren van het ontwikkelproces.

T - Taak: Een iteratief en systematisch proces volgen voor het ontwikkelen van het leeftijdsschatting model, met continue evaluatie en verbetering.

A - Actie: We hebben het CRISP-DM framework toegepast en zijn iteratief de stappen van Business Understanding tot uiteindelijk Deployment aan het volgen. Version control gebeurt via GitLab waar alle experimenten, model checkpoints en resultaten worden bijgehouden. We hebben systematisch geëxperimenteerd met verschillende hyperparameters, architectuur aanpassingen, en data augmentation strategieën, zoals ook eerder is besproken.

R - Resultaat: De technische review beoordeelde dit als "Boven verwachting": "Jullie hebben een iteratie uitgevoerd, wat genoeg is voor dit punt in het project. Jullie evalueren de werking niet alleen door te kijken naar metriecken, maar ook door naar de diepere betekenis van deze waarden te kijken in verhouding tot de context." (Bewijs: BEW-007) We hebben een werkend model met gedocumenteerde experimenten en kunnen de evolutie van het model reproduceren via Git history.

R - Reflectie: Mijn ervaring met CRISP-DM tijdens mijn afstudeerstage heeft baat gehad tijdens dit project. Version control via Git was ook een vanzelfsprekende keuze die soepel kwam door de eerdere ervaring met Git-based systemen. De feedback dat we "naar de diepere betekenis kijken in verhouding tot context" laat

zien dat de praktische relevantie van de context net zo belangrijk als de technische details is. Een beperking is dat we pas één iteratie hebben gedaan op dit punt in het blok. Voor een volledig systematisch proces zouden meerdere iteraties ideaal zijn. Dit is deels een tijds kwestie, maar ook een keuze om breed te werken (alle competenties) versus diep te gaan (veel iteraties op één aspect).

T - Transfer: Tijdens dit blok is het nodig om nog meerdere iteraties van modellen te creëren, en vergelijksmoellen voor classification te maken (zoals beschreven bij A4). Voor het eindrapport is het essentieel dat elke iteratie goed gedocumenteerd staat met duidelijke motivatie waarom we bepaalde richting(en) kozen.

C. EVALUEREN EN MONITOREN

C1: Maatschappelijke impact evalueren

S - Situatie: Dit is een competentie waar ik nog weinig concrete ervaring mee heb in AI-projecten. Het dichtstbijzijnde voorbeeld was tijdens het TNO Speaker Recognition project, waar ik opmerkte dat ons model stiltes tussen zinnen koppelde aan specifieke sprekers, waardoor gesprekken met veel stiltes altijd aan dezelfde spreker werden toegewezen. We behandelden dit voornamelijk als een technische kwestie en zijn niet diepgaand ingegaan op de bredere implicaties voor individu en maatschappij. Dit is een gebied waar ik graag meer systematische ervaring in wil ontwikkelen, vooral het proces van het identificeren en bespreken van mogelijke ongewenste consequenties van AI-oplossingen.

T - Taak: De maatschappelijke impact van leeftijdscontrole AI evalueren, ongewenste consequenties identificeren, en hierover in gesprek gaan met vakgenoten.

A - Actie: Ik heb de error flow ontwikkeld die verschillende foutscenario's en edge cases doorloopt. Hierin staan alle "ongewenste" situaties beschreven waar een dergelijk systeem mee om moet kunnen gaan. (Bewijs: BEW-003)

R - Resultaat: De error flow documenteert systematisch verschillende ongewenste consequenties, zoals:

- System errors: Kritieke systeemfout. Mitigatie: terug naar volledig menselijke verificatie.
- Spoofing attempts: Ondernijning systeem geloofwaardigheid. Mitigatie: anomaly detection met directe escalatie naar employee.
- Retry logic: Juist positioneren van het gezicht totdat er een heldere foto gemaakt kan worden.

De flow heeft als basis gediend voor discussies binnen het team en met docenten over hoe het systeem moet reageren op verschillende scenario's.

R - Reflectie: Het ontwerpen van de error flow dwingt je om expliciet na te denken over wat er misgaat in plaats van alleen de happy flow. Dit zorgt ervoor dat je vooraf nadenkt over "ongewenste" omstandigheden en hoe je daarop reageert. De flow helpt om elk scenario te doorlopen en concrete mitigaties te bedenken (retry logic, fallbacks, etc...). Een beperking is dat ik vooral technische mitigaties heb bedacht maar minder over bijvoorbeeld foute voorspellingen heb nagedacht tijdens het ontwerpen.

T - Transfer: Voor het eindrapport ga ik de error flow uitbreiden met een vervolgstappen als er een foute voorspelling gemaakt wordt, of als de medewerker verdacht gedrag ziet.

Voor toekomstige projecten neem ik mee: begin met error flows en happy flows. Dit forceert om kritisch te denken over wat er mis kan gaan en hoe je daar op moet reageren.

C2: Kwaliteitscriteria toepassen

S - Situatie: Ik heb ervaring met standaard evaluatiemetriecken zoals accuracy, precision, en recall in praktisch elk AI-project. Mijn afstudeerproject ging verder door ook resource demand, explainability, en scalability te analyseren als onderdeel van het totaalplaatje. Echter, aspecten zoals robustness en model complexity zijn weinig tot niet geanalyseerd in mijn eerdere projecten. Over deze kwaliteitscriteria wil ik graag meer leren en toepassen, vooral in de context van production AI-systems waar deze factoren belangrijk kunnen zijn.

T - Taak: Het leeftijdsschatting model evalueren aan de hand van erkende kwaliteitscriteria zoals robustness, performance, scalability, explainability, model complexity, en resource demand.

A - Actie: Naast standaard metrics (MAE) heb ik de business critical zones gedefinieerd die specifieke risico-percentages berekenen voor verschillende leeftijdsgroepen. Deze zones zijn gevisualiseerd in scatterplots die de voorspelling versus werkelijke leeftijd tonen met gekleurde regio's voor verschillende risiconiveaus. (Bewijs: BEW-005)

R - Resultaat: De technische review beoordeelde dit als "Op niveau": "Evaluatie is goed uitgevoerd voor de belangrijkste metrics. De Business-Critical Zone Analysis voegt veel waarde toe en maakt inzichtelijk hoe het model presteert voor de belangrijkste groep." (Bewijs: BEW-007) Bij de individuele opdracht echter, bleek uit de M&T feedback dat robustness, scalability, explainability, model complexity, en resource demand onvoldoende zijn geadresseerd. Er ontbreekt een confusion matrix en sommige keuzes zijn niet gefundeerd onderbouwd. (Bewijs: BEW-008)

R - Reflectie: Het pragmatiseren van metriecken is intuïtief makkelijker te voltooien voor mij dan vaak het kwantificeren van abstracte metriecken in het kader van robustness, explainability, en model complexity. Robustness (wat als de verlichting slecht is?), explainability (waarom voorspelt het model deze leeftijd?), model complexity (hoe definieer je een model als "complex"?) zijn allemaal relevante vragen die niet of nauwelijks zijn geadresseerd. Het gebrek van een confusion matrix was helaas een stomme fout: ik was vergeten de confusion matrix te plotten. De plot functie had ik er namelijk tijdelijk even uitgehaald en vergeten terug te plaatsen.

T - Transfer: Voor volgende M&T opdrachten en het eindrapport zal het belangrijk zijn om alle type metriecken langs te gaan en aan te duiden, in plaats van het houden bij de basis metriecken.

C3: Prototype testing met stakeholders

S - Situatie: Deze competentie heb ik meerdere keren toegepast tijdens mijn afstudeerproject, waar we iteratief de churn prediction modellen op ongeziene data hebben getest en de resultaten analyseerden met stakeholders. Het proces omvatte het presenteren van bevindingen aan CTO en CEO, het toepassen van hun feedback, en het aanpassen van de aanpak gebaseerd op business requirements. Deze cyclus van testen, evalueren, en aanpassen was belangrijk voor het ontwikkelen van een praktisch bruikbare oplossing. Ik ben sterk in het vertalen van technische resultaten naar business impact en het faciliteren van feedback loops met stakeholders.

T - Taak: Samen met stakeholders het prototype voor het leeftijdscontrole project analyseren met behulp van prototypes en inzichten, feedback meenemen in het proces.

A - Actie: Ik heb een eigen paper prototype ontwikkeld in Figma dat verschillende screens toont voor de happy flow en error flow. Dit prototype bevat een basis representatie van de applicatie om de flow te laten zien. Na bespreking met het team hebben we één gemeenschappelijk paper prototype gemaakt dat elementen van verschillende ontwerpen combineert. (Bewijs: BEW-006)

R - Resultaat: Het gemeenschappelijke paper prototype is gebruikt tijdens de feedback sessie met de CMD-studenten. Het prototype visualiseert zowel de happy flow (succesvolle leeftijdsverificatie) als verschillende error scenarios. Op basis van informele feedback tijdens deze review en groepsdiscussies hebben we enkele aanpassingen gedaan aan de UI, zoals duidelijkere instructies voor het correct positioneren van het gezicht, een explicietere feedback loop wanneer het systeem geen leeftijd kan schatten, en meer inzicht in o.a. privacy-regulaties via een info knop. Het prototype vormde de basis voor ons uiteindelijke implementatie en hielp om vroegtijdig discussies te voeren over edge cases en gebruikersinteractie.

R - Reflectie: Het maken van een paper prototype voordat we aan implementatie begonnen bleek waardevol. Het maakte abstract denken over flows concreet. Discussies over "wat gebeurt er als..." werden makkelijker te voeren met visuele ondersteuning.

T - Transfer: Voor de rest van het blok plannen we om voor de expo onze interface te testen binnen de groep en medestudenten. Specifiek willen we testen:

- Begrijpen ze de interface en instructies zonder uitleg?
- Voelt de interactie natuurlijk aan binnen hun bestaande kassaworkflow?
- Hoe ervaren ze de balans tussen automatisering en menselijke controle? Het laatste punt zal zeer interessant zijn om te zien van medestudenten die een hele andere mate van automatisering hebben toegepast.

D. ZELFSTURING

D1: Feedback verzamelen en verwerken

S - Situatie: Dit is een competentie waar ik nog geen concrete voorbeelden van heb binnen specifiek AI-projecten. Hoewel ik wel ervaring heb met het ontvangen en verwerken van feedback in algemene projectcontext (zoals tijdens mijn afstudeerstage), heb ik dit nog niet systematisch gedocumenteerd of gestructureerd toegepast in de context van mijn eigen ontwikkeling binnen AI. Dit is een belangrijke competentie die ik graag wil ontwikkelen door een meer bewuste en gedocumenteerde aanpak van feedbackverwerking in mijn leerproces.

T - Taak: Actief feedback verzamelen over eigen kennis, vaardigheden, en ontwikkeling, en deze feedback gebruiken om werk te verbeteren door blok 1 heen.

A - Actie: Ik houd dit persoonlijk logboek bij waarin ik feedback van coaches, docenten (tech workshops, M&T, ontwerp workshops, filosofie & ethiek), en medestudenten (peer feedback) documenteer. Voor elke feedback noteer ik wat de feedback was, waar het over ging, hoe het aansluit bij mijn leerdoelen, en wat ik ermee doe. (Bewijs: dit logboek zelf)

R - Resultaat: Het voortgangsgesprek en de STARRT-reflectie zijn het directe bewijs van systematisch feedback verzamelen en verwerken. Ik heb o.a. feedback ontvangen op:

- Technische review (groep): meerdere competenties, vooral B en C (Bewijs: BEW-007)

- M&T Opdracht 1a (individueel): vooral A2, B1, B2, C2 (Bewijs: BEW-008)
- PvA Artikel (individueel): vooral D2, D3 (Bewijs: BEW-009)

Deze feedback heb ik verwerkt in bovenstaande reflecties en heb concrete acties geformuleerd voor verbetering (zie Transfer secties per competentie).

R - Reflectie: In mijn bachelor en afstudeerstage kreeg ik ook feedback maar documenteerde ik het niet zo expliciet. Ik ben bekend met de STARR-methode, maar niet de STARRT-methode, en ook net tot deze mate. Ook het ontvangen van feedback is een nieuw onderdeel voor mij binnen de context van competentieaanduidingen. Ik merk dat ik sommige feedback makkelijker accepteer dan andere. Technische feedback (bijv. "je data leakage is fout") is makkelijk te accepteren en wil ik meteen oplossen. Feedback over softere aspecten (bijv. "je onderbouwing is summier") vind ik moeilijker te verwerken omdat het minder concreet voelt. Als de feedback abstracter aanvoelt is het soms lastiger te verwerken, maar uiteraard net zo waardevol.

T - Transfer: Voor de rest van het blok ga ik:

- Feedback gelijk integreren in de volgende iteratie.
- Een feedback-ronde houden met mijn groep.

Voor toekomstige projecten is het handig om feedbackverwerking een gewoonte te maken, en het gelijk aan te pakken en er op te reflecteren.

D2: Onderzoek opzetten en rapporteren

S - Situatie: Deze competentie heb ik sterk ontwikkeld tijdens mijn afstudeerstage door het schrijven van een research article over mijn churn prediction systeem. Het onderzoek volgde academische standaarden en ik communiceerde de resultaten zowel schriftelijk (via het artikel) als mondeling met mijn CTO en stagebegeleider. Het artikel is momenteel in preprint fase, wat aantoont dat de kwaliteit op academisch niveau is. Ik ben comfortabel met het opzetten van onderzoek, het volgen van methodologieën, en het rapporteren volgens academische standaarden. Mijn sterke punt ligt in het maken van de vertaalslag tussen technische bevindingen en praktische implicaties.

T - Taak: Onderzoek opzetten, uitvoeren, en rapporteren voor artikel binnen het leeftijdscontrole project volgens academische standaarden.

A - Actie: Ik heb een plan van aanpak geschreven met onderzoeksvraag, literatuuronderzoek, en methodologie. De onderzoeksvraag is: "What dataset characteristics are most critical for achieving both high accuracy and fairness in facial age estimation systems?" Het artikel volgt de IEEE template en bevat literatuuronderzoek naar bias in facial recognition, dataset vergelijkingen, en een voorgestelde experimentele methodologie. (Bewijs: BEW-009)

R - Resultaat: De PvA feedback beoordeelde D2 als "Op niveau" maar D3 als "Hier is nog werk voor nodig". Specifieke feedback punten: (Bewijs: BEW-009) Positief:

- Probleemstelling volgt logisch uit literatuur over bias en retail automation
- Research question is beknopt en helder
- Sterke literatuurreview met meerdere bronnen

Verbeterpunten:

- Planning mist specifieke datums/weken per milestone
- Doelstelling is te algemeen: wie is de doelgroep? Academisch of praktijk?
- Research question is breed en vraagt eigenlijk meerdere sub-vragen
- Methodologie te beperkt:
 - Geen train-test split strategie
 - Geen preprocessing details
 - Onduidelijk wanneer/waarom VGG16 wordt gebruikt
 - Geen training en validation protocol
- Bronnen: let op kwaliteit van white papers (3 van Yoti)

R - Reflectie: Dit is een competentie waar ik concrete ervaring mee heb uit mijn afstudeerstage (research article), maar de feedback laat zien dat ik hier nog groei in heb zitten. Mijn afstudeerstage artikel was vooral descriptive (wat hebben we gebouwd?), terwijl een master artikel analytical moet zijn (wat zijn de diepere inzichten?). De opmerkingen zijn ieder valide. Door de drukte van het groepsproject had ik helaas minder tijd voor individuele opdrachten, en dat laat zien dat de resultaten er onder lijden. Betere planning en structuur kan hierbij helpen.

T - Transfer: Voor het finale artikel ga ik:

- Research question opsplitsen in main question + 2-3 sub-questions voor focus
- Doelgroep expliciet maken: schrijf voor academisch publiek (dataset researchers)
- Methodologie uitbreiden:
 - Volledige train/val/test split strategie
 - Preprocessing pipeline in detail
 - Training protocol (epochs, learning rate, optimizer, etc.)
 - Validation protocol (metrics, wanneer stoppen, hoe best model selecteren)
- Literatuur review doen op white papers: vervang niet-peer-reviewed bronnen waar mogelijk
- Planning toevoegen met concrete weken per milestone

Deze stappen zouden mijn artikel op niveau moeten brengen.

D3: Robuuste onderzoeksmethoden kiezen

S - Situatie: Dit heb ik toegepast tijdens mijn afstudeerstage via het gebruik van het CRISP-DM framework, uitgebreid literatuuronderzoek, en het naleven van MDPI publicatiestandaarden voor mijn research article. Het onderzoek bevatte systematische vergelijking van verschillende machine learning technieken, temporal validation strategies, en evaluatiemethodologieën. Ik heb ervaring met het selecteren van passende methodologieën gebaseerd op onderzoeksvragen en het onderbouwen van methodologische keuzes met literatuur. Mijn aanpak is systematisch en evidence-based, wat heeft geresulteerd in een bijdrage aan de kennis in het vakgebied van churn prediction voor subscription-based businesses.

T - Taak: Robuuste en valide onderzoeksmethoden kiezen voor het beeldverwerking project en bijdragen aan ontwikkeling van nieuwe kennis in AI-vakgebied.

A - Actie: Ik pas het CRISP-DM framework toe op het leeftijdsvoorspellingsproject:

- Business Understanding: Stakeholder interviews (AH medewerker) voor requirements en context
- Data Understanding: Preventief dataset onderzoek, EDA op UTKFace en IMDB-Cleaned
- Data Preparation: Preprocessing pipeline, data augmentation strategieën

- Modeling: VGG16-gebaseerd model ontwikkeling met systematische experimenten
- Evaluation: Business critical zones als evaluatiemethodiek

Daarnaast voer ik literatuuronderzoek uit naar moderne facial age estimation methodiek en onderbouw methodologische keuzes met academische bronnen voor het artikel.

R - Resultaat:

Het CRISP-DM framework helpt om het project systematisch aan te pakken en iteraties te structureren. Echter, de PvA feedback beoordeelde D3 als "Hier is nog werk voor nodig": de methodologie is "te beperkt" met ontbrekende details over essentiële delen van een standaard methodologie, zoals train-test splits, preprocessing, en training/validation protocols. (Bewijs: BEW-009)

R - Reflectie:

Ik dacht dat ik robuuste methoden gebruikte (CRISP-DM, literatuur, systematische experimenten), maar de feedback laat zien dat een methodologie toepassen niet vanzelfsprekend is. Hoewel is bij het leeftijdsvoorspellingsproject door de verschillende stappen ben gegaan, komen kleine methodologische fouten nog wel omhoog.

T - Transfer:

Voor het artikel ga ik:

- Train/test split strategie expliciet beschrijven (stratificatie op leeftijd én etniciteit)
- Preprocessing pipeline volledig documenteren (face detection, alignment, normalization, augmentation)
- Training/validation protocol specificeren (epochs, learning rate, optimizer, early stopping criteria)