

00:00:00.000 --> 00:00:01.060

Anouk Wolters
Once it gets here.

00:00:09.468 --> 00:00:10.138

Machine Learning Engineer/Projectmanager
OK.

00:00:10.628 --> 00:00:17.678

Anouk Wolters
Hoi, dat zou jij je rol en verantwoordelijkheden in de NBA case kunnen beschrijven.

00:00:20.218 --> 00:00:23.568

Machine Learning Engineer/Projectmanager
Ja, mijn huidige en en historische rol.

00:00:25.008 --> 00:00:32.098

Anouk Wolters
Ja wacht, ik zit even te denken, Misschien is het handig Als ik eerst nog even een korte recap doe van waar ik mee bezig ben en wat het doel is.

00:00:32.588 --> 00:00:33.708

Machine Learning Engineer/Projectmanager
Ja, Dat is goed.

00:00:34.888 --> 00:00:42.298

Anouk Wolters
Nou, Ik ben dus nu bezig met mijn thesis en hiervoor onderzoek naar de ontwikkeling en het gebruik van machine learning.

00:00:42.888 --> 00:00:52.838

Anouk Wolters
Als socio technische systemen en dat betekent dat ik niet Alleen kijk naar de modellen zelf, maar ook hoe zo'n model geïntegreerd wordt in een sociale en organisatorische context.

00:00:53.468 --> 00:00:53.738

Machine Learning Engineer/Projectmanager
Ja.

00:00:53.478 --> 00:01:09.128

Anouk Wolters
En in dit interview wil ik er graag achter komen hoe het ontwikkelen van een machine learning use case In de praktijk in zijn werk gaat. En Ik wil graag kijken in hoeverre jullie en hoe jullie omgaan met de sociotechnische dimensies die Ik heb geïdentificeerd In de literatuur.

00:01:09.898 --> 00:01:23.948

Anouk Wolters

En uiteindelijk werk ik toe naar het ontwerpen van een methode of proces waarmee dit sociotechnische perspectief kan worden ingebracht In de machine learning die verder ben praktijk om zo bij te dragen aan een veilig en verantwoord gebruik van machine learning.

00:01:09.928 --> 00:01:10.348

Machine Learning Engineer/Projectmanager
Ja.

00:01:24.658 --> 00:01:25.058

Machine Learning Engineer/Projectmanager
Ja.

00:01:25.628 --> 00:01:27.778

Anouk Wolters
Dus dat een beetje als context?

00:01:27.718 --> 00:01:29.558

Machine Learning Engineer/Projectmanager
Duidelijk? Ja, duidelijk.

00:01:29.548 --> 00:01:38.398

Anouk Wolters
Ja, dus dan, Als we terug kunnen gaan naar je rol en verantwoordelijkheden. Inderdaad, hoe dat ook is, veranderd over de tijd ook wel handig om mee te nemen.

00:01:39.098 --> 00:01:39.458

Machine Learning Engineer/Projectmanager
Ja.

00:01:41.108 --> 00:02:08.458

Machine Learning Engineer/Projectmanager
Ja, Ik ben bij de NBA engine wat de naam is voor de machine, eigenlijk de engine waar machine learning gebruikt wordt in eerste instantie verantwoordelijk geweest als machine learning engineer, dus zowel voor het ontwikkelen van een model Als het ontwerp daarvan, dus het analyseren van de data hebben we gedaan. Wat gaan we gebruiken, hoe gaan we die data gebruiken? Ontwikkeling daarvan en later overgestapt naar een rol waarin ik meer als projectmanager fungeerde en een andere machine engineer ontwikkeling over heeft genomen.

00:02:09.178 --> 00:02:14.018

Machine Learning Engineer/Projectmanager
daar ook de modellen nog heeft aangepast en verbeteringen doorgevoerd.

00:02:14.728 --> 00:02:15.098

Anouk Wolters
Oké.

00:02:15.878 --> 00:02:20.388

Anouk Wolters

En, zijn we iets kunnen vertellen over hoe het begin van dit project eruit zag?

00:02:21.868 --> 00:02:45.398

Machine Learning Engineer/Projectmanager

Ja zeker, dus We hebben in het begin een soort designfase zijn we ingegaan, of eigenlijk een potentie analyse noemen we dat om te kijken of of het het waard was om hier machine learning te gaan toepassen. Daarbij hebben we met een aantal business stakeholders gezeten, waaronder een segment marketeer, dus iemand die In het segment.

00:02:48.098 --> 00:03:01.088

Machine Learning Engineer/Projectmanager

Retail business retail Ik weet niet hoe het heet Retail Investment zat geloof ik dus in ieder geval ging over retail beleggings rekeningen van klanten bij [de bank], een data analist en hun manager.

00:03:02.268 --> 00:03:22.878

Machine Learning Engineer/Projectmanager

En met hen Samen hebben we een soort design fase ingegaan waarin wij gekeken hebben. Inderdaad daar en tegelijkertijd interviews bij hen hebben gehouden en dat Samen hebben gebracht en zegt, oké, nou, volgens ons is deze business case te halen. Kunnen we het zo implementeren? Gaat ons dat zolang kosten en en moet dat ook zo in zijn werk gaan, dus moet het zo straks geïntegreerd worden In de.

00:03:23.538 --> 00:03:24.668

Machine Learning Engineer/Projectmanager

In de organisatie.

00:03:26.448 --> 00:03:55.938

Machine Learning Engineer/Projectmanager

Dus Dat was de design fase. Daarna hebben we een implementatiefase ingeschoten waarin we in uit mijn hoofd 3 maanden het geïntegreerd hebben. Dus ontwikkeld en geïntegreerd in hun bestaande tooling. En daarna hebben we nog een kleine fase gedaan waarin we de operationalisatie gedaan hebben. Dus We hebben de data analyst als operator aangewezen en daarbij ook de documentatie goed vastgelegd en gezorgd dat deze persoon ook kon valideren dat alles goed ging in deze engine dat ie deed wat hij moest doen en ook enige mate van controle had.

00:03:55.988 --> 00:03:57.678

Machine Learning Engineer/Projectmanager

Door wat configuraties aan te passen.

00:03:58.858 --> 00:03:59.218

Anouk Wolters

OK.

00:04:00.948 --> 00:04:08.178

Anouk Wolters

En, Waarom is dit model of deze engine? In de eerste instantie ontwikkeld? Wat is het doel ervan?

00:04:10.518 --> 00:04:12.218

Machine Learning Engineer/Projectmanager

Ja, het doel is.

00:04:14.018 --> 00:04:44.918

Machine Learning Engineer/Projectmanager

Het het beter informeren van Van retail klanten van [de bank], dus klanten, particuliere klanten met een met een spaarrekening of een betaalrekening over wat beleggen voor hen kan betekenen en en de hoop. Daarmee is het doel is dat meer klanten ook gaan beleggen, Omdat dat nou ja zo vaak voor klanten blijkt een een goede keuze te zijn om met hun vermogen om te gaan, maar ook voor [de bank] zelf een interessante optie is.

00:04:45.408 --> 00:04:51.838

Machine Learning Engineer/Projectmanager

Dus de conversie van klanten Zonder beleggingsrekening naar klanten die beleggen bij [de bank].

00:04:52.868 --> 00:04:53.278

Anouk Wolters

Oké.

00:04:53.998 --> 00:05:01.628

Anouk Wolters

En voordat Er is begonnen aan die ontwikkelings fase, zijn er dan bepaalde criteria of requirements die jullie opstellen?

00:05:05.438 --> 00:05:25.408

Machine Learning Engineer/Projectmanager

Ja zeker dus allereerst commercieel hè? Moet er een bepaalde businesscase zijn, dus Het gaat In het geval e-mail uitingen, dus We willen klanten gaan benaderen met informatieve e-mails over beleggingen, dus Je moet weten dat je genoeg klanten kunt benaderen überhaupt en dat je hè? Met met een inschatting van een bepaalde conversie daar ook genoeg.

00:05:26.348 --> 00:05:30.458

Machine Learning Engineer/Projectmanager

Een goed genoeg businesscase uit kunt halen om überhaupt die machine te ontwikkelen.

00:05:32.708 --> 00:05:34.858

Machine Learning Engineer/Projectmanager

Dus Dat is er eentje. Daarnaast hebben we ook gezegd dat we.

00:05:36.178 --> 00:05:48.208

Machine Learning Engineer/Projectmanager

Een vink willen hebben van hun compliance officer, dus ook de data die we gebruiken. We gebruiken in principe persoonlijke data van [de bank], hoewel het niet om NAW weging, maar wel over hun producten.

00:05:49.728 --> 00:05:53.718

Machine Learning Engineer/Projectmanager

Dat we dat ook op een verantwoorde manier gebruiken en dat er iemand daar ook naar gekeken heeft en dat goedgekeurd heeft.

00:05:58.418 --> 00:06:03.048

Machine Learning Engineer/Projectmanager

Ik denk ja, dus dat zijn de twee grootste denk ik dus commercieel en compliant, dat zijn de grootste

00:06:03.848 --> 00:06:04.248

Anouk Wolters

Ja.

00:06:05.108 --> 00:06:12.418

Anouk Wolters

En en wie zijn er dan betrokken bij het opstellen van de deze requirements?

00:06:15.078 --> 00:06:29.068

Machine Learning Engineer/Projectmanager

Nou het team wat ik net noemde, dus de de de manager van marketing intelligence en data analyst, een segment marketeer dus meer als business expert en en en ik.

00:06:29.788 --> 00:06:30.048

Anouk Wolters

Ja.

00:06:31.548 --> 00:06:38.338

Anouk Wolters

Oké en hoe wordt het dan uiteindelijk geëvalueerd? Of het systeem ook voldoet aan deze eisen?

00:06:40.728 --> 00:06:44.968

Machine Learning Engineer/Projectmanager

Nou bij de compliance is dat heel simpel, dat moet en dat moet een vinkje hebben, dus We hebben.

00:06:47.138 --> 00:06:56.118

Machine Learning Engineer/Projectmanager

Laten zien welke we konden aangeven welke data we gingen gebruiken voor welk doel. En daar heeft iemand toen opgezegd akkoord. Dit kun je zo gaan gebruiken.

00:06:57.538 --> 00:07:19.288

Machine Learning Engineer/Projectmanager

En voor de commercie hebben we periodieke evaluaties gedaan van de conversieratio's en gekeken ook ja hoe dat hoe dat ging en daarbij hè? Dus een soort tests ingericht om te kijken om zo objectief mogelijk de impact van de NBA engine te meten. Dus je weet nooit zeker of Mensen beleggingsrekening openen Omdat jij die mail heb gestuurd. Ja of nee?

00:07:20.118 --> 00:07:40.668

Machine Learning Engineer/Projectmanager

Dus We hebben ook gekeken naar. We hebben ook Mensen die de NBA engine niet zou mailen uit zichzelf hebben we ook een mail gestuurd om te kijken wat dan het effect was. En We hebben ook Mensen die de NBA wel wilde mailen hebben geen mail gestuurd om te kijken wat daar de conversie was en doorr eigenlijk die die twee controlegroepen naast de de groep te zetten die we wel gemailld hadden, konden we zien wat de verhoogde conversie was.

00:07:42.388 --> 00:07:42.728

Anouk Wolters

Oké.

00:07:44.638 --> 00:07:45.148

Anouk Wolters

En.

00:07:46.258 --> 00:07:51.408

Anouk Wolters

Waren er ook bepaalde kwaliteitsvoorwaarden die bestaan binnen de bank waar jullie rekening mee moesten houden.

00:07:55.168 --> 00:07:58.258

Machine Learning Engineer/Projectmanager

Ja zeker dus [de bank] vindt het heel belangrijk om.

00:07:59.328 --> 00:08:08.958

Machine Learning Engineer/Projectmanager

Goed met goed met data om te gaan en ook echt gewoon hè? Ver binnen de lijnen van Van wat wat mag binnen GDPR te kleuren dus We hebben.

00:08:09.668 --> 00:08:26.838

Machine Learning Engineer/Projectmanager

We hebben bewust voor gekozen, bijvoorbeeld om om niet te kijken naar geslacht, Omdat we ook hè uit eerdere analyses binnen [de bank] gezien hebben dat mannen bijvoorbeeld vaker hoofd rekening houder zouden zijn zijn dan vrouwen percentueel binnen [de bank] dus.

00:08:27.888 --> 00:08:33.388

Machine Learning Engineer/Projectmanager

Dat zou betekenen dat hè, je kans zegt hebt dat mannen vaak geïnformeerd worden over wat beleggen doet.

00:08:35.258 --> 00:08:43.478

Machine Learning Engineer/Projectmanager

We hebben ook besloten om niet publieke data over bijvoorbeeld postcodes te gebruiken, dus heb ik niet naar PC4 gekeken. Dat hebben we achterwege gelaten.

00:08:44.428 --> 00:08:48.358

Anouk Wolters

En, hoe bepaal je dan wat wat je wel en niet meeneemt?

00:08:49.818 --> 00:08:50.798

Anouk Wolters

Waar baseer je dat op?

00:08:54.948 --> 00:08:56.828

Machine Learning Engineer/Projectmanager

Op op persoonlijke intuïtie.

00:08:58.278 --> 00:08:58.548

Machine Learning Engineer/Projectmanager

Ja.

00:08:59.518 --> 00:09:01.628

Machine Learning Engineer/Projectmanager

dus we hebben zelf gekeken naar wat zouden mogelijkerwijs.

00:09:03.478 --> 00:09:19.348

Machine Learning Engineer/Projectmanager

Wat is dan direct gevoelige data, hè? Dus bij nou ja, bij geslacht of bij nou postcode ook postcode kan Natuurlijk kan gevoelig zijn, Omdat postcode ook een proxy kan zijn voor bepaalde etniciteit in een wijk zitten of zoiets dus dus ja dat.

00:09:18.868 --> 00:09:19.198

Anouk Wolters

Ja.

00:09:20.858 --> 00:09:30.438

Anouk Wolters

Oké en uiteindelijk heb je dus bepaalde data niet meegenomen. Denk je dat er alsnog een mogelijkheid zou zijn dat er toch bias in sluipt?

00:09:33.358 --> 00:09:40.938

Machine Learning Engineer/Projectmanager

Nou, ik durf niet hard te claimen dat dat absoluut niet zo is, Maar ik, Ik denk dat dat meevalt, want Ik denk echt dat we vrij. Ja.

00:09:43.628 --> 00:09:47.098

Machine Learning Engineer/Projectmanager

Ja daar nauwelijks gegevens over over meenemen.

00:09:48.178 --> 00:09:48.578

Machine Learning Engineer/Projectmanager
Ja.

00:09:50.408 --> 00:10:00.998

Anouk Wolters

En ook en tijdens het ontwikkelen van het model de use case. Hoe speelt dan de context waarin het model uiteindelijk gebruikt gaan worden? Een rol in die ontwikkeling?

00:10:04.388 --> 00:10:06.138

Machine Learning Engineer/Projectmanager
De organisatorische context, bedoel je?

00:10:06.918 --> 00:10:07.268

Anouk Wolters

Ja.

00:10:09.598 --> 00:10:10.248

Machine Learning Engineer/Projectmanager
Nou.

00:10:12.778 --> 00:10:13.538

Machine Learning Engineer/Projectmanager
Even kijken.

00:10:19.338 --> 00:10:20.858

Machine Learning Engineer/Projectmanager
Er zijn Er zijn, denk ik.

00:10:22.828 --> 00:10:25.748

Machine Learning Engineer/Projectmanager
Er zijn twee dingen belangrijk, dus wij kiezen.

00:10:27.068 --> 00:10:57.778

Machine Learning Engineer/Projectmanager

Allereerst, hoe ga je zorgen dat de organisatie en zeker de Mensen die ermee moeten werken ook begrijpen wat er gebeurt? Dus We hebben bijvoorbeeld, We hebben niet voor de aller moeilijkste modellen of de allermoeilijkste features gekozen om daar als concreet voorbeeld voor te geven. We hebben een set een set aan data, dus een subset en de data die een persoon omschrijft en dat zijn heel concreet. Heeft deze personen app geïnstalleerd van [de bank]? Hoeveel rekeningen heeft?

00:10:57.828 --> 00:11:29.078

Machine Learning Engineer/Projectmanager

Deze persoon, hoe vaak loggen ze in op internet bankieren? Die hebben allemaal samengevoegd tot één feature, Omdat ze min of meer hetzelfde zeiden. En We hebben die feature uitgelegd aan [de bank] als dit is de betrokkenheid dus eigenlijk laat deze feature zien hoe betrokken klant is bij [de bank] en door dat in een ding te verpakken voordat je dat

In het model stopt was het ook heel makkelijk voor voor die data analist om dat uit te dragen naar naar de rest van de organisatie. Dus hij kon eigenlijk zeggen, van ja, we kijken naar een stuk of 7 features 7 datapunten en eentje daarvan is de betrokkenheid.

00:11:30.368 --> 00:11:31.208

Anouk Wolters

OK ja.

00:11:30.958 --> 00:12:01.508

Machine Learning Engineer/Projectmanager

En de en de tweede is is de de controle die ze over hebben, dus dus meer het het verpakken in software waarbij ze kunnen zien wat er gebeurt, bijvoorbeeld hoeveel Mensen worden geselecteerd. Hoe hoog is de de kans hè? Dus het model? We hebben een bepaalde kans gegeven dat dat model denkt dat iemand converteert en die en die kans ja die die lag tussen de 0% en de 100%. Ja, we wisten niet precies of dat of of 80% echt 80% was, Maar het gaf wel meer aan dat het model dacht ja vrij hoog.

00:12:05.358 --> 00:12:15.418

Anouk Wolters

En even kijken, je had het net al een beetje over data die jullie gebruiken. Hoe en door wie is de data verzameld?

00:12:20.008 --> 00:12:21.308

Machine Learning Engineer/Projectmanager

De data is.

00:12:25.018 --> 00:12:54.038

Machine Learning Engineer/Projectmanager

Het is een subset van de data die [de bank] in zijn data warehouse heeft staan, dus [de bank] heeft op een centrale plek uit meerdere systemen data verzameld en die data daar hebben wij een subset van beschikbaar gekregen als marketing intelligence. En daar hebben wij weer een subset van gepakt en in het verladen van die data van het data warehouse naar dit systeem naar het subsysteem. Daar hebben we bewust gekozen om bijvoorbeeld geen NAW mee te nemen.

00:12:54.098 --> 00:12:56.088

Machine Learning Engineer/Projectmanager

Dus die data, hadden we letterlijk niet.

00:12:57.838 --> 00:12:58.328

Machine Learning Engineer/Projectmanager

Ja.

00:12:59.498 --> 00:13:03.618

Anouk Wolters

En hoe wordt een goede datakwaliteit bewerkstelligd?

00:13:04.698 --> 00:13:06.458

Anouk Wolters

En wat is een goede datakwaliteit?

00:13:12.638 --> 00:13:14.198

Machine Learning Engineer/Projectmanager

Nou dus enerzijds

00:13:18.548 --> 00:13:46.218

Machine Learning Engineer/Projectmanager

Ja, dus datakwaliteit is is hoe schoon is die data? Hoe compleet is het heb je het van Van alle Mensen beschikbaar en ik moet zeggen dat daar daar zijn Natuurlijk al heel veel stappen voor gebeurd voordat het in zo'n datawarehouse beland, dus daar hebben we op zich al hele schone data gekregen. Ik weet Alleen dat wat niet zo betrouwbaar was is precies wanneer iemand klant was geworden. Dat was een veld wat wat niet heel betrouwbaar was, dus dus gewoon qua qua schoonheid.

00:13:46.898 --> 00:14:07.508

Machine Learning Engineer/Projectmanager

Zag het er goed uit. De tweede is een representeert het goed wat je wil doen en en en daar moesten we af en toe creatief zijn. Dus Ik weet bijvoorbeeld dat we wilden graag het beursklimaat meenemen, dus economisch klimaat door naar de beurs te kijken. Maar we kunnen geen externe data zomaar aanspreken bij [de bank], want dat zijn hele interne beveiligde netwerken.

00:14:09.138 --> 00:14:22.688

Machine Learning Engineer/Projectmanager

Dus hebben we de fonds koersen van hun eigen fondsen die ze wel intern hadden hebben we als een soort proxy gepakt. Ja, en Dat is Natuurlijk niet perfect, want die koersen dat zijn hele hele groene sustainable koersen en die bewegen anders dan de beurs zelf.

00:14:24.128 --> 00:14:24.508

Anouk Wolters

Ja.

00:14:25.568 --> 00:14:31.498

Anouk Wolters

Heeft dat dan ook. Uiteindelijk denk je invloed gehad op de model hoe het model werkt en de output daarvan.

00:14:36.588 --> 00:14:39.258

Machine Learning Engineer/Projectmanager

Ja, Ik denk dat dat uiteindelijk wel.

00:14:42.618 --> 00:14:51.328

Machine Learning Engineer/Projectmanager

Ja moeilijk te moeilijk, moeilijk te zeggen, weet ik niet. Ik weet niet of Als het echt tot de

echte beurs features hadden, die hebben gewoon niet hebben, ook nooit naast kunnen leggen. Dus dat dat weet ik niet.

00:14:55.038 --> 00:15:03.278

Anouk Wolters

Wat voor risico's of fouten verder kunnen ontstaan tijdens het ontwikkelen van het machine learning systeem.

00:15:08.108 --> 00:15:09.338

Machine Learning Engineer/Projectmanager

Tijdens het ontwikkelen.

00:15:14.298 --> 00:15:15.558

Machine Learning Engineer/Projectmanager

Ja, kijk, Je kunt.

00:15:23.158 --> 00:15:31.598

Machine Learning Engineer/Projectmanager

Nou één van de dingen, Er is iets fout gegaan, dus een van de dingen die mis is gegaan, is dat we.

00:15:32.938 --> 00:15:36.968

Machine Learning Engineer/Projectmanager

Het is een heel concreet voorbeeld, Ik weet niet of het zo gedetailleerd moet, maar.

00:15:40.278 --> 00:16:00.408

Machine Learning Engineer/Projectmanager

We hebben dus die. We hebben dus een feature meegeven over de fonds koers, hè? Om om het beursklimaat representeren. En Dat is eigenlijk een getal ongeveer hoeveel hoeveel die gestegen was over een week en over een maand dus we hadden gezegd, nou, wekelijks is Hij is hij zo hard omhoog gaan, zo hard omlaag gaan of procentueel en over een maand is het deze zodat je een beetje korte en lange termijn had.

00:16:01.128 --> 00:16:27.458

Machine Learning Engineer/Projectmanager

Maar de combinatie van die twee cijfers was eigenlijk een unieke key. in de data set dus elke week waarop je op data keek was er maar één een key dus dus eigenlijk kon daarmee het model onthouden over welke week de data ging. En heeft hij Daarom In de In de development van het eerste model op geoverfit. Dus dan wist hij bijvoorbeeld ja in.

00:16:29.978 --> 00:16:46.158

Machine Learning Engineer/Projectmanager

om dat te vertalen dus in mei krijgen mensen vakantiegeld dus dan openen veel Mensen een beleggingsrekening en dat en dat model kon dus aan de hand van soort van beursklimaat onthouden. Ah dit is de maand mei of dus mei 2020. Hier waren veel Mensen geconverteerd, dus dat dat zal wel weer zou zijn.

00:16:46.848 --> 00:16:50.378

Anouk Wolters

Nou, ja oke, en hoe ga je dan om met?

00:16:51.258 --> 00:16:55.218

Anouk Wolters

Fouten, die er in zouden kunnen sluipen. Of hoe voorkom je die?

00:16:57.058 --> 00:16:59.788

Anouk Wolters

Nou Misschien eerst, hoe voorkom je dat er fouten insluipen?

00:17:07.638 --> 00:17:11.398

Machine Learning Engineer/Projectmanager

Ja even kijken, hoe voorkom je dat?

00:17:17.678 --> 00:17:18.228

Machine Learning Engineer/Projectmanager

Ja.

00:17:19.478 --> 00:17:50.938

Machine Learning Engineer/Projectmanager

hoe zouden we dit nou voorkomen hebben? Ja, ik. Ik denk uiteindelijk dat we het hadden kunnen voorkomen als we nog een iets breder test scenario hadden. Dus We hebben nu hebben Natuurlijk. We hebben data gebruikt om te trainen. Een hele grote subset en we hebben het ook nog op een klein periodetje getest en daarin zaten net niet de uitzonderingsgevallen die ervoor zorgde dat het model zich raar gedragen. Dus dat kwam gewoon. We hadden maar 1,5 jaar aan data dus en moesten het grootste deel voor het trainen gebruiken, maar Als we daar ook nog 1,5 jaar op hadden kunnen simuleren hoe dat dingen had gedaan voordat hem echt.

00:17:50.988 --> 00:17:53.018

Machine Learning Engineer/Projectmanager

Live zetten, dan waren we er eerder achter gekomen.

00:17:53.618 --> 00:17:53.948

Anouk Wolters

Ja.

00:17:56.428 --> 00:17:57.358

Anouk Wolters

En Misschien.

00:17:58.098 --> 00:18:01.478

Anouk Wolters

Voor andere situaties zijn er bepaalde.

00:18:02.608 --> 00:18:05.268

Anouk Wolters

Dingen jullie die je dan checkt, of?

00:18:06.938 --> 00:18:12.578

Anouk Wolters

Bepaalde strategie om te voorkomen dat voordat een model live gaat.

00:18:13.488 --> 00:18:15.088

Anouk Wolters

Fout erin zijn gesloten.

00:18:19.088 --> 00:18:25.808

Machine Learning Engineer/Projectmanager

Ja, dus, Er zijn zeker, Er zijn zeker manieren om dat om dat te checken, dus wat we.

00:18:28.308 --> 00:18:44.118

Machine Learning Engineer/Projectmanager

Ja, Dat is wat we eigenlijk gedaan hebben om dat goed in te kunnen schatten dus we hebben simulaties gedraaid, wel dus over dus wat ik eerder zei wat we uitgebreider hadden kunnen doen, Maar we hebben gesimuleerd. Oké, hoeveel hè gegeven dit model? Hoeveel mails zou je wekelijks nou versturen en en zitten daar gekke dingen tussen, zien we wel gekke dingen.

00:18:46.218 --> 00:19:05.088

Machine Learning Engineer/Projectmanager

Daarnaast hè? Heel intuïtief kijken we ook. Dus oké, ik verwacht dat als de fonds koersen keihard omhoog gaan, dat meer Mensen ook een beleggingsrekening willen openen, dus dat die er nog meer mails wil versturen, is dat ook zo. Dus Als ik nu Als ik nu zelf een beetje aan de knoppen gaat draaien en en doe alsof de fonds koersen omhoog gaan, zie ik dat dan ook of niet?

00:19:05.698 --> 00:19:05.988

Anouk Wolters

Ja.

00:19:08.328 --> 00:19:11.548

Machine Learning Engineer/Projectmanager

Ja dus dus intuïtief empirisch denk ik.

00:19:12.118 --> 00:19:12.388

Anouk Wolters

Ja.

00:19:13.038 --> 00:19:13.338

Anouk Wolters

Oké.

00:19:15.888 --> 00:19:23.058

Anouk Wolters

En stel, er komt dan toch nog In de model output fout of bias? Hoe ga je daarmee om?

00:19:22.318 --> 00:19:22.738

Machine Learning Engineer/Projectmanager

Ja.

00:19:27.008 --> 00:19:29.698

Machine Learning Engineer/Projectmanager

Nou, Ik denk in eerste instantie aan.

00:19:30.598 --> 00:19:45.418

Machine Learning Engineer/Projectmanager

Nou kijk of het een kritieke fout is of niet, of je denkt dat het een hele gevoelige fout is. In ons geval was het geen hele gevoelige fout. Dat betekent gewoon dat hij teveel mails verstuurde in een week waarin dat niet zou moeten, dus een beetje ja, niet het juiste moment pakte om sommige Mensen te mailen.

00:19:46.678 --> 00:19:58.448

Machine Learning Engineer/Projectmanager

Dus daar hebben we toen een soort hotfix voor gemaakt en gezegd, nou, dit lossen we even tijdelijk zo op door een stukje code om dat model heen te zetten. En Als we straks met een V2 bezig gaan, dan gaan we dit oplossen.

00:19:59.098 --> 00:20:01.658

Machine Learning Engineer/Projectmanager

En toen in die V2 heeft [andere ML engineer] dat ook opgelost.

00:20:02.388 --> 00:20:02.768

Anouk Wolters

Ja.

00:20:03.448 --> 00:20:07.058

Machine Learning Engineer/Projectmanager

Als dat een kritieke fout was geweest, ja, dan waren we natuurlijk niet gaan draaien.

00:20:08.598 --> 00:20:09.348

Anouk Wolters

Nee precies.

00:20:10.148 --> 00:20:10.518

Machine Learning Engineer/Projectmanager

Ja.

00:20:11.028 --> 00:20:11.538

Anouk Wolters

En als.

00:20:13.238 --> 00:20:16.418

Anouk Wolters

Stel, Het is toch gaan draaien in?

00:20:17.218 --> 00:20:18.428

Anouk Wolters

productie en dan.

00:20:19.258 --> 00:20:21.498

Anouk Wolters

Komt er een foute model output?

00:20:21.898 --> 00:20:22.428

Machine Learning Engineer/Projectmanager

Ja.

00:20:22.878 --> 00:20:23.708

Anouk Wolters

Hoe ga je daarmee om?

00:20:26.148 --> 00:20:42.708

Machine Learning Engineer/Projectmanager

Nou, We hebben in dit geval. We hadden een Human in the loop, dus de data analist was onze achtervang daarin. Hij draaide het model, keek naar de output en als hij het goed vond, dan zette hij het pas door naar het mailprogramma, dus Dat is Dat was de eerste achtervang.

00:20:44.488 --> 00:20:57.958

Machine Learning Engineer/Projectmanager

En, dat hebben we ook. We hebben ook echt de eerste paar mail badges die dat ding wilde doen, ook echt Samen door gegraven, dus is het nou goed, zitten er geen Mensen in die vorige week al gemaïld zijn gewoon goed gecontroleerd.

00:21:00.788 --> 00:21:03.598

Machine Learning Engineer/Projectmanager

Ja en Als het toch In het productie landt, als het toch misgaat.

00:21:06.438 --> 00:21:09.908

Machine Learning Engineer/Projectmanager

Ja nog, Dan hadden hem niet gebruikt, dan hadden we hem uitgezet denk ik.

00:21:10.758 --> 00:21:14.718

Machine Learning Engineer/Projectmanager

Of of wil je weten wat we gedaan hadden Als we het hadden moeten corrigeren?

00:21:15.098 --> 00:21:16.918

Anouk Wolters

Ja dat ook nog.

00:21:17.638 --> 00:21:18.438
Machine Learning Engineer/Projectmanager
En.

00:21:22.478 --> 00:21:25.088
Machine Learning Engineer/Projectmanager
Ja, dus ik denk In het In het.

00:21:27.448 --> 00:21:28.788
Machine Learning Engineer/Projectmanager
In het ergste geval.

00:21:29.958 --> 00:21:33.808
Machine Learning Engineer/Projectmanager
Als we dan hadden we Mensen meerdere keren gemaild.

00:21:34.648 --> 00:21:59.938
Machine Learning Engineer/Projectmanager
Ja dan dan hadden we daar denk ik. Daar hadden we dan niet nog een excuus mail achteraan gestuurd of iets dan hadden we hem gewoon uit de lucht gehaald en dan hadden we het probleem opgelost. Het is, Het is geen, Het is geen kritieke toepassing eigenlijk, dus We hebben dus een menselijke controle, maar ook menselijke achtervang. Dus Als het Als het Als het niet goed was, dan kon hem gewoon uitzetten en en ging die persoon zelf mail sturen en dan selecteer je die 500 willekeurige Mensen.

00:21:42.448 --> 00:21:42.728
Anouk Wolters
Ja.

00:22:00.758 --> 00:22:01.048
Anouk Wolters
Ja.

00:22:02.148 --> 00:22:02.498
Anouk Wolters
Oké.

00:22:03.528 --> 00:22:04.228
Anouk Wolters
En.

00:22:06.738 --> 00:22:07.708
Anouk Wolters
Even kijken.

00:22:10.738 --> 00:22:24.568
Anouk Wolters
Als nu het model dan in productie gebruikt wordt, zouden er dan ook zou je kunnen denken

aan bepaalde veranderingen externe factoren die invloed kunnen hebben op de performance van het model over tijd?

00:22:26.688 --> 00:22:27.658

Machine Learning Engineer/Projectmanager
Ja zeker.

00:22:29.788 --> 00:22:30.518

Anouk Wolters
Gegeven.

00:22:31.328 --> 00:22:32.778

Machine Learning Engineer/Projectmanager
Ja dus dus dus zijn?

00:22:34.388 --> 00:22:38.608

Machine Learning Engineer/Projectmanager
Intern zijn er een hoop dingen die zouden kunnen veranderen, hè? Dus zij kunnen.

00:22:40.378 --> 00:22:57.728

Machine Learning Engineer/Projectmanager
producten kunnen bijvoorbeeld ophouden te bestaan. Stel je voor zeg voordat [de bank] geen betaalrekeningen meer wil doen, maar alleen maar spaarrekeningen, dan valt die data weg of internetbankieren houdt ermee op of de app houden ze mee op, of dat wordt een ander veld of zoiets. Dat kan allemaal.

00:23:00.528 --> 00:23:03.328

Machine Learning Engineer/Projectmanager
Een andere is meer externe invloeden, Dat is, Dat is.

00:23:06.168 --> 00:23:14.018

Machine Learning Engineer/Projectmanager
Bijvoorbeeld corona is zo'n ding dus ten tijde van corona, hè? Dus in maart 2020 dipte die beurs enorm.

00:23:15.738 --> 00:23:27.188

Machine Learning Engineer/Projectmanager
En zo'n dip van de beurzen, Dat is, Dat is niet een scenario. Dat is echt een crash. Dat is niet een scenario waar zo'n model op getraind is en zelfs al had het een keer gezien, had het nog niet goed gereageerd, waarschijnlijk, want dat is één keer.

00:23:29.108 --> 00:23:44.938

Machine Learning Engineer/Projectmanager
Dus in in dat soort scenario's en die kun je van tevoren meten, hè? Dus je weet een soort van de bandbreedte waarbinnen de beurs mag bewegen, hoe snel dat fluctueert enzo en als dat te veel is zoals in maart 2020. Echt een wereldwijde economische crash dan.

00:23:45.618 --> 00:23:51.048

Machine Learning Engineer/Projectmanager

Op de financiële markt, dan zou je, dan zou je niet op dat model moeten kunnen varen.

00:23:52.078 --> 00:23:58.098

Anouk Wolters

Oké, en hoe ga je dan om met deze mogelijkheid voor veranderingen In het gebruik van het model?

00:23:59.398 --> 00:24:14.328

Machine Learning Engineer/Projectmanager

Naar ik, ja, Ik denk niet dat je een model kunt maken wat daar wat voor zo'n voor zulke uitzonderingssituaties goed reageert of dat het de moeite waard is om dat te doen. Dus ik zou in dat geval gewoon zeggen. Daar moeten Mensen even weer de hand op het stuur hebben.

00:24:15.428 --> 00:24:22.868

Anouk Wolters

Nou ja, en hoe wordt dat dan ingericht? Dat dat er wel In de gaten wordt gehouden of er niet veranderingen zijn die invloed hebben?

00:24:26.038 --> 00:24:30.338

Machine Learning Engineer/Projectmanager

Ja, dus, We hebben nu in een in een in een later stadium nog.

00:24:31.598 --> 00:24:47.608

Machine Learning Engineer/Projectmanager

checks ingebouwd dus Er zijn Er zijn checks inderdaad die dit soort checks doen kijken of de data goed is en en of een bepaalde aannames waar zijn. En als dat niet zo is dan, dan draait hij niet en dan wordt er een mail gestuurd naar de naar de genen die er verantwoordelijk voor is van hé, let op, Er zijn gekke dingen aan de hand.

00:24:48.278 --> 00:24:50.998

Anouk Wolters

Dat dan checks In de input data.

00:24:51.418 --> 00:24:51.718

Machine Learning Engineer/Projectmanager

Ja.

00:24:52.708 --> 00:24:53.038

Machine Learning Engineer/Projectmanager

Ja.

00:24:54.088 --> 00:25:02.578

Anouk Wolters

En zijn er nog andere dingen die gemonitord worden aan het model en wat wordt er dan gewoon door?

00:25:10.268 --> 00:25:14.478

Machine Learning Engineer/Projectmanager

Ja dus de even kijken dus dus de aan de input kant worden gemonitord.

00:25:15.928 --> 00:25:16.838

Machine Learning Engineer/Projectmanager

Aan de output.

00:25:17.548 --> 00:25:32.608

Machine Learning Engineer/Projectmanager

Ja niet automatisch denk ik, dus worden we checken niet of het, dus We hebben wel een bepaalde verwachting. Hoeveel mails die mag sturen. Dus in een slechte week mogen dat er nou mogen dat er bijna nul zijn, maar In de goede week mogen dat wel maximaal tweeduizend zijn als dat in één keer twintigduizend waren geweest, dan is er echt iets.

00:25:34.198 --> 00:25:35.788

Machine Learning Engineer/Projectmanager

Dus dat, ja, dus dat wordt.

00:25:36.528 --> 00:25:39.938

Machine Learning Engineer/Projectmanager

Nu menselijk gecontroleerd hebben we niet expliciet hchecks op.

00:25:40.598 --> 00:25:48.748

Machine Learning Engineer/Projectmanager

En verder en technisch, dus we kijken gewoon of het of het allemaal draait of geen bugs in zitten en geen andere issues ontstaan.

00:25:49.898 --> 00:25:54.358

Anouk Wolters

Oké en wat als er wel een issue in staat, hoe werkt dat dan?

00:25:55.538 --> 00:26:06.888

Machine Learning Engineer/Projectmanager

Ja, Het is, dan hebben we dan is. De eerste lijn is dus die data analyst die operator is. Die krijgt een mail met het is niet goed gegaan. Die gaat dan kijken of hij dat zelf op kan lossen.

00:26:08.088 --> 00:26:13.458

Machine Learning Engineer/Projectmanager

Niet het geval is, dan komt hij bij ons uit en dan gaan wij er naar kijken dus dat is een tweede lijn support.

00:26:11.928 --> 00:26:12.168

Anouk Wolters

Ja.

00:26:13.908 --> 00:26:14.228

Anouk Wolters

OK.

00:26:14.848 --> 00:26:17.368

Anouk Wolters

En hoe is bepaald wat er gemonitord moet worden?

00:26:22.128 --> 00:26:29.488

Machine Learning Engineer/Projectmanager

Nou ja, deels door door onze kennis denk ik, dus We hebben voor een aantal suggesties voorgedragen.

00:26:30.188 --> 00:26:39.938

Machine Learning Engineer/Projectmanager

En deels doordat we weten welke fouten door kunnen ontstaan, dus het datawarehouse wil wel eens te laat zijn data verladen, waardoor wij niet een nieuwe week op kunnen halen.

00:26:40.648 --> 00:26:46.078

Machine Learning Engineer/Projectmanager

Ja, Dat is 3 keer gebeurd en daardoor wisten we ook dat we daarop moesten checken en dat dat dat ie daarmee om moest gaan.

00:26:46.938 --> 00:26:47.818

Anouk Wolters

Ja oké.

00:26:48.658 --> 00:26:58.358

Anouk Wolters

En is er ook sprake van een naast monitoring van een andere soort evaluatie gedurende dat systeem wordt gebruikt.

00:27:05.378 --> 00:27:14.648

Machine Learning Engineer/Projectmanager

Nou van het systeem zelf niet van de resultaten wel, dus we kijken zelf ook periodiek naar zijn de resultaten wat je ervan verwacht gaan niet omlaag.

00:27:15.618 --> 00:27:15.898

Anouk Wolters

Ja.

00:27:16.008 --> 00:27:22.038

Machine Learning Engineer/Projectmanager

Maar en, Maar dat is niet. Wat doen we? Dat doen we ook nog handmatig. Niet zo dat hij een automatische monitoring op die feedback loop zit ofzo.

00:27:22.548 --> 00:27:22.988

Anouk Wolters
OK.

00:27:22.778 --> 00:27:23.318

Machine Learning Engineer/Projectmanager
Dat nog niet.

00:27:24.008 --> 00:27:26.988

Anouk Wolters
En, hoe vaak wordt dat gecheckt geëvalueerd?

00:27:29.148 --> 00:27:38.548

Machine Learning Engineer/Projectmanager
Dat weet ik niet, want dat doet met name die operator tegenwoordig toevallig volgende week maandag hebben er weer eentje is. Wij hebben het nu al 4 5 keer gedaan, denk ik.

00:27:39.718 --> 00:27:40.558

Anouk Wolters
En, hoe lang is dat?

00:27:42.158 --> 00:27:42.898

Machine Learning Engineer/Projectmanager
1,5 jaar.

00:27:43.388 --> 00:27:44.458

Anouk Wolters
Oké ja.

00:27:49.678 --> 00:27:50.528

Anouk Wolters
Even kijken.

00:27:51.178 --> 00:27:54.718

Anouk Wolters
En zijn er dan bepaalde punten waarop je evalueert of hoe?

00:27:57.528 --> 00:27:59.598

Anouk Wolters
wordt dat vormgegeven?

00:28:00.348 --> 00:28:01.148

Machine Learning Engineer/Projectmanager
Ja, dus je?

00:28:01.948 --> 00:28:05.138

Machine Learning Engineer/Projectmanager
Eigenlijk de punten waarop je ook van tevoren je.

00:28:07.698 --> 00:28:37.208

Machine Learning Engineer/Projectmanager

De business case op heb gemaakt dus hoeveel hoe hoog is de conversie ten opzichte van die controlegroepen. Hoeveel geld hebben Mensen in gelegd op die beleggingsrekening? Want dat bepaalt uiteindelijk ook de de marge die je erop pakt. Hoeveel Mensen in totaal benaderd hebt, hè? Zijn er ook, zijn er Mensen die altijd overslaat is, is een restbakje. En zo ja, wie zijn dat dan? En Waarom slaat die die over? Dus er zit ook een stukje explainability bij, dus We hebben ook voorgelegd in eerste instantie van Hey.

00:28:37.398 --> 00:28:38.528

Machine Learning Engineer/Projectmanager

Misschien wil je wel?

00:28:39.388 --> 00:28:40.658

Machine Learning Engineer/Projectmanager

Een stukje logica die?

00:28:44.088 --> 00:28:50.638

Machine Learning Engineer/Projectmanager

Random Mensen selecteert die nooit geselecteerd worden om te kijken of we daar toch niet iets missen die in één keer wel.

00:28:51.318 --> 00:29:03.828

Machine Learning Engineer/Projectmanager

Blijken heel goed reageren op die mails, ook al denken wij van niet, maar daar hebben ze zegt dat ze dat niet willen. Ze willen liever eerst gewoon dan analyseren Waarom die Mensen nooit geselecteerd worden en wat daar achter zit. En daar volgens.

00:29:05.008 --> 00:29:05.258

Anouk Wolters

Ja.

00:29:05.288 --> 00:29:06.438

Machine Learning Engineer/Projectmanager

Misschien later alsnog iets mee doen.

00:29:06.998 --> 00:29:09.098

Anouk Wolters

Zijn daar al inzichten op.

00:29:10.368 --> 00:29:11.178

Machine Learning Engineer/Projectmanager

Niet dat Ik weet.

00:29:11.938 --> 00:29:12.378

Machine Learning Engineer/Projectmanager

Nee.

00:29:13.708 --> 00:29:14.158

Anouk Wolters
Oké.

00:29:15.528 --> 00:29:21.318

Anouk Wolters
Even kijken, hoe wordt de model output gepresenteerd aan de operator?

00:29:24.608 --> 00:29:34.998

Machine Learning Engineer/Projectmanager
In principe door een een lijst te genereren met daarin de personen, dus de de relatie nummers die een mail kunnen krijgen.

00:29:36.948 --> 00:29:45.718

Machine Learning Engineer/Projectmanager
Waarbij we eerst ook nog hadden gezegd Waarom? Dus is deze door de door het model geselecteerd of is dit een testgroep bijvoorbeeld een randomized control?

00:29:48.628 --> 00:29:54.768

Machine Learning Engineer/Projectmanager
En daarnaast hebben we ook wel eens die voorspellingskans, dus de de de kans.

00:29:56.088 --> 00:30:01.558

Machine Learning Engineer/Projectmanager
Ja, dus de kans dat we denken dat iemand converteert hoe hoog die is, die Laten we ook zien.

00:30:02.718 --> 00:30:03.778

Anouk Wolters
Oké en.

00:30:03.278 --> 00:30:06.088

Machine Learning Engineer/Projectmanager
Maar dat is meer In de evaluaties dan dan In de dag dagelijks.

00:30:07.388 --> 00:30:07.928

Machine Learning Engineer/Projectmanager
Stappen.

00:30:08.258 --> 00:30:08.918

Anouk Wolters
Ja OK.

00:30:09.718 --> 00:30:11.028

Anouk Wolters
En en.

00:30:11.678 --> 00:30:21.608

Anouk Wolters

Hoe kan de operator die input van of eigenlijk de output van het model gebruiken, om te uiteindelijk te beslissen of er wel of geen?

00:30:22.338 --> 00:30:24.268

Anouk Wolters

Mail wordt gestuurd, hoe doet hij dat?

00:30:25.998 --> 00:30:31.528

Machine Learning Engineer/Projectmanager

En die operator kan zien In het Mailprogramma Selligent gebruiken zij.

00:30:33.248 --> 00:30:41.918

Machine Learning Engineer/Projectmanager

Kan zo'n operator zien hebben we deze persoon al eerder gemailld en zo ja, wanneer en en hij kan ook steekproefsgewijs in onze database kijken van hé.

00:30:42.568 --> 00:30:46.008

Machine Learning Engineer/Projectmanager

Wie is deze persoon? Vind ik het ook logisch dat deze persoon gemailld wordt?

00:30:48.538 --> 00:30:50.918

Machine Learning Engineer/Projectmanager

Die data kunnen wij ook, daar hebben wij ook.

00:30:52.808 --> 00:30:58.218

Machine Learning Engineer/Projectmanager

Wat voor klaar zegt dat hij makkelijk kan kijken van hé, wat heeft die engine? Waarom heeft die engine dat geselecteerd?

00:31:01.028 --> 00:31:09.878

Machine Learning Engineer/Projectmanager

En Als ie dat ok vindt en en je merkte eigenlijk dat die aan het begin heel veel controleerde daar heel strak op zat. En nu dat eigenlijk bijna niet meer doet omdat ie het wel vertrouwdt.

00:31:10.658 --> 00:31:10.958

Anouk Wolters

Oké.

00:31:12.658 --> 00:31:16.668

Machine Learning Engineer/Projectmanager

Nu dus die die relatie nummers doorzet naar een mailprogramma zegt, hier gaat deze mailen.

00:31:17.628 --> 00:31:17.968

Anouk Wolters

Ja.

00:31:18.718 --> 00:31:25.748

Anouk Wolters

En je zei al dat hier ook een stukje explainability voor wordt gebruikt. Kun je dat wat meer toelichten? Hoe dat eruit ziet?

00:31:29.518 --> 00:31:30.868

Machine Learning Engineer/Projectmanager

Ja, dus, We hebben niet.

00:31:32.448 --> 00:31:38.668

Machine Learning Engineer/Projectmanager

We gebruiken niet expliciet explainability zoals dus We gebruiken Geen explainers om voor elke.

00:31:42.928 --> 00:32:02.248

Machine Learning Engineer/Projectmanager

Om om voor elke persoon te Laten zien wat zwaar mee gewogen is. We gebruiken we Laten eigenlijk input features zien en hebben die zo geselecteerd dat hij ook begrijpbaar zijn voor ze, zodat ook deze persoon dit profiel hebben in een model gestopt en op basis daarvan is deze score uitgekomen.

00:32:03.948 --> 00:32:06.028

Anouk Wolters

Dus, eigenlijk is het gewoon.

00:32:07.268 --> 00:32:08.748

Anouk Wolters

Totaal interpreteerbaar, of niet?.

00:32:10.188 --> 00:32:12.148

Anouk Wolters

Wat mist hij

00:32:14.598 --> 00:32:18.588

Anouk Wolters

Wat kan hij wel en wat kan hij niet interpreteren uit het model?

00:32:18.988 --> 00:32:22.418

Machine Learning Engineer/Projectmanager

Nou, Ik denk dat hij het nog best lastig vind dan echt te interpreteren. Waarom?

00:32:23.978 --> 00:32:27.578

Machine Learning Engineer/Projectmanager

Waarom iemand geselecteerd is dus?

00:32:27.348 --> 00:32:27.628

Anouk Wolters

Ja.

00:32:29.998 --> 00:32:30.508
Machine Learning Engineer/Projectmanager
Ja.

00:32:33.408 --> 00:32:35.948
Anouk Wolters
Oke dus hij heeft wel iets van inzicht, maar.

00:32:36.588 --> 00:32:37.388
Anouk Wolters
Hij zou niet.

00:32:38.558 --> 00:32:46.048
Anouk Wolters
Dus stel een klant vraagt, Waarom ben ik geselecteerd? Hiervoor? Zou hij dat dan kunnen uitleggen op basis van.

00:32:48.128 --> 00:32:49.598
Anouk Wolters
De inrichting van het Model en output.

00:32:50.298 --> 00:32:52.038
Machine Learning Engineer/Projectmanager
Ik denk dat 'een' uitleg kan geven.

00:32:53.238 --> 00:32:53.558
Anouk Wolters
Oké.

00:32:55.468 --> 00:32:55.788
Anouk Wolters
Ja.

00:32:59.298 --> 00:33:06.618
Anouk Wolters
Even kijken je zei net al dat hij in eerste instantie heel veel controleerde, en dat hij het wat meer vertrouwt.

00:33:04.598 --> 00:33:05.028
Machine Learning Engineer/Projectmanager
Ja.

00:33:07.278 --> 00:33:13.098
Anouk Wolters
Hoe ga je om met het gedrag van de operator ten aanzien van het model?

00:33:21.138 --> 00:33:27.898
Anouk Wolters

Wordt daar In het ontwikkelingsproces ook rekening mee gehouden dat gedrag ook een invloed kan hebben op het uiteindelijke gebruik.

00:33:30.748 --> 00:33:31.338

Machine Learning Engineer/Projectmanager
Nou.

00:33:33.778 --> 00:33:35.348

Machine Learning Engineer/Projectmanager
Ja, dus, We hebben ja.

00:33:36.688 --> 00:33:45.888

Machine Learning Engineer/Projectmanager
Dus We hebben hem geprobeerd controle te geven over het model door configuraties, dus We hebben een plek waar hij.

00:33:46.718 --> 00:33:48.778

Machine Learning Engineer/Projectmanager
Dingen aan kan passen, bijvoorbeeld hoe?

00:33:50.638 --> 00:34:02.908

Machine Learning Engineer/Projectmanager
Ja hoe makkelijk het model komt, selecteert hè dus dus dus Als het een goede week is, moeten we dan 1000 of tweeduizend Mensen gemaild worden? Daar kan hij mee spelen Als hij die aantallen te hoog te veel vindt, fluctueren of te weinig.

00:34:04.608 --> 00:34:11.518

Machine Learning Engineer/Projectmanager
Hij kan hem ook uitzetten, dus die en die draait automatisch. Elke week doet hij een analyse die kan hij uitzetten, daar zijn knopje voor.

00:34:12.618 --> 00:34:13.958

Machine Learning Engineer/Projectmanager
We hebben tegelijkertijd.

00:34:14.788 --> 00:34:17.388

Machine Learning Engineer/Projectmanager
Ook gezegd tegen hem, dus een beetje.

00:34:20.708 --> 00:34:42.728

Machine Learning Engineer/Projectmanager
Ja geprobeerd te begeleiden in hoe je dat gebruikt van ga niet teveel veranderen, want Als je ook elke week zit te spelen met die output, dan is straks niet meer te evalueren Of of of het nou effect heeft of niet wat je wat je doet. Dus probeer daar ook gewoon vooral 3 maanden niet aan te zitten en dan eens te kijken of dat of dat je bevalt en het dan weer eens aan te passen en niet elke week dat ding omhoog of omlaag.

00:34:44.658 --> 00:34:45.148
Machine Learning Engineer/Projectmanager
Ja.

00:34:45.928 --> 00:34:46.308
Anouk Wolters
Ja.

00:34:46.938 --> 00:34:53.048
Machine Learning Engineer/Projectmanager
En, Dat is eigenlijk de enige manier waarop hij denk ik in zijn gedrag invloed uit kan oefenen op wat dat ding op dat model doet.

00:34:54.228 --> 00:35:02.318
Anouk Wolters
Ja precies en denk je. Ook wordt er ook rekening mee gehouden dat Misschien wat hij denkt of hoe hij ermee omgaat.

00:35:04.848 --> 00:35:06.978
Anouk Wolters
Zorgt dat de model output Misschien?

00:35:07.638 --> 00:35:10.178
Anouk Wolters
Minder goed wordt gebruikt, dan zou kunnen?

00:35:12.478 --> 00:35:13.498
Machine Learning Engineer/Projectmanager
Nee? Dat denk ik niet.

00:35:12.788 --> 00:35:14.488
Anouk Wolters
Dat hij biased is?

00:35:15.648 --> 00:35:20.488
Machine Learning Engineer/Projectmanager
Nee, Ik denk dat Hij heeft met name een controlerende rol, dus Ik denk dat dat wel meevalt.

00:35:21.918 --> 00:35:22.238
Anouk Wolters
OK.

00:35:24.428 --> 00:35:25.238
Anouk Wolters
Even kijken.

00:35:32.268 --> 00:35:41.638
Anouk Wolters

In hoeverre zijn jullie als externe partij verantwoordelijk voor het model en de output daarvan?

00:35:44.778 --> 00:35:50.278

Machine Learning Engineer/Projectmanager

We zijn verantwoordelijk voor de technische ontwikkeling, maar niet voor het gebruik of de resultaten.

00:35:51.798 --> 00:35:52.248

Anouk Wolters

Oké.

00:35:53.028 --> 00:35:59.048

Anouk Wolters

En stel het wordt al gebruikt dus in productie en blijkt dat er iets in de.

00:35:59.968 --> 00:36:03.948

Anouk Wolters

In het in het model zelf niet goed is.

00:36:04.928 --> 00:36:06.408

Anouk Wolters

Waardoor de output.

00:36:07.358 --> 00:36:10.058

Anouk Wolters

Ook niet goed is. Hoe zit dat dan?

00:36:08.958 --> 00:36:09.378

Machine Learning Engineer/Projectmanager

Ja.

00:36:14.478 --> 00:36:15.558

Machine Learning Engineer/Projectmanager

Ja dus dan.

00:36:16.298 --> 00:36:17.168

Machine Learning Engineer/Projectmanager

Dan zullen wij.

00:36:18.978 --> 00:36:25.728

Machine Learning Engineer/Projectmanager

We hebben geen service overeenkomst met [de bank] dus we zullen ze helpen om dat te fizen en dat hebben we ook altijd gedaan.

00:36:27.138 --> 00:36:32.848

Machine Learning Engineer/Projectmanager

Maar we zijn niet verantwoordelijk voor die voor die output of ook niet voor foute output, nee.

00:36:33.268 --> 00:36:33.548

Anouk Wolters

Ja.

00:36:33.938 --> 00:36:44.808

Machine Learning Engineer/Projectmanager

Nee, We hebben ook wel duidelijk gemaakt dat ze daar dat het hun verantwoordelijkheid is, hè, dus dat ze zelf ook zelf moeten aangeven dat ze er mee eens zijn, manier waarop de data gebruikt wordt en en ja, dat dat.

00:36:47.428 --> 00:36:48.428

Machine Learning Engineer/Projectmanager

wat het model doet.

00:36:49.608 --> 00:36:50.768

Anouk Wolters

Ja ja oké.

00:36:54.598 --> 00:36:55.108

Anouk Wolters

Even kijken.

00:37:03.368 --> 00:37:14.428

Anouk Wolters

Ja tijdens het de ontwikkeling van deze use cases of Misschien In het algemeen use cases hebben jullie dan bepaalde processen of workflows die jullie volgen?

00:37:16.288 --> 00:37:17.068

Machine Learning Engineer/Projectmanager

Ja zeker.

00:37:18.128 --> 00:37:19.558

Anouk Wolters

Zou je daar wat over vertellen?

00:37:18.308 --> 00:37:18.798

Machine Learning Engineer/Projectmanager

Dus ik.

00:37:20.678 --> 00:37:26.568

Machine Learning Engineer/Projectmanager

Nou, Ik denk wat heel klassiek is aan de manier waarop dit gedaan hebben is. Het is het.

00:37:27.198 --> 00:37:43.498

Machine Learning Engineer/Projectmanager

Het analyseren om te kijken of dit een use case is die het waard is hè dus echt, nou is er een business case. Heb je de data? Hoe ligt dat erbij? Kunnen we uit de business expert ook horen we ook dat we kwalitatief denken dat daar iets te halen valt.

00:37:45.898 --> 00:38:06.548

Machine Learning Engineer/Projectmanager

En In de ontwikkeling ook of na de ontwikkeling eigenlijk. Sorry, dus dan heb je de ontwerpfase, dan komt die ontwikkeling en dan na die ontwikkeling heb je. Natuurlijk hebben we altijd validatie periode, dus waarin we gewoon gaan draaien en kijken, is dit wat We willen ja of nee? En om er een beetje vertrouwen in te krijgen en daarna het operationaliseren daarvan dus de juiste Mensen.

00:38:07.278 --> 00:38:11.638

Machine Learning Engineer/Projectmanager

op de juiste processen inrichten om te zorgen dat dat met minimale overhead gewoon gebruikt kan worden.

00:38:13.448 --> 00:38:15.858

Machine Learning Engineer/Projectmanager

Op een goede manier, Dat is denk ik vrij standaard.

00:38:16.908 --> 00:38:17.288

Anouk Wolters

Ja.

00:38:17.918 --> 00:38:32.328

Anouk Wolters

En worden dan In het begin. In de eerste fase wordt daar dan ook rekening mee gehouden met bepaalde risico's die het met zich meebrengt om zo'n model te ontwikkelen? Of ja?

00:38:33.018 --> 00:38:38.758

Anouk Wolters

Kwetsbaarheden die zouden kunnen optreden waar tijdens de ontwikkeling en het gebruik rekening mee moet worden gehouden.

00:38:43.678 --> 00:38:46.548

Machine Learning Engineer/Projectmanager

Nou, ja, dus Ik denk dat ik.

00:38:47.448 --> 00:39:06.208

Machine Learning Engineer/Projectmanager

Ja wat ik eerder noemde, dus we hebben geprobeerd de bias er zoveel mogelijk uit te halen door niet naar NAW of naar geslacht of naar postcode te kijken. We hebben ook gewaarschuwd dat er uitzonderingssituaties zijn waarin je dat maar model niet wil gebruiken zoals beurs, crashes of juist. Ja, een hele rare uitzonderingssituaties.

00:39:08.238 --> 00:39:09.178

Machine Learning Engineer/Projectmanager

Ja dat denk ik.

00:39:09.648 --> 00:39:13.088

Anouk Wolters

En dat dingen die je tijdens het ontwikkelen van een use case.

00:39:14.958 --> 00:39:19.168

Anouk Wolters

Kijkt in welk geval wat belangrijk is. Of hebben jullie bijvoorbeeld ook een soort?

00:39:21.308 --> 00:39:27.548

Anouk Wolters

Checklist of iets van dit zijn dingen die je waar je altijd op moet letten voordat je een model gaat ontwikkelen.

00:39:28.648 --> 00:39:37.458

Machine Learning Engineer/Projectmanager

Nee, dat hebben we niet. Nee, dat hangt heel erg van de use case af. Ik zou deze ook niet terug kunnen vinden in een andere use case die ik gedaan heb of zo, Dat is altijd anders.

00:39:36.788 --> 00:39:37.028

Anouk Wolters

Nee.

00:39:38.058 --> 00:39:38.338

Anouk Wolters

Ja.

00:39:39.128 --> 00:39:40.208

Anouk Wolters

Er zijn er wel.

00:39:41.548 --> 00:39:43.868

Anouk Wolters

Even Kijken dus bij [external ML engineering/consulting firm] in het algemeen

00:39:44.498 --> 00:39:53.718

Anouk Wolters

Dan gebruiken jullie die wat je net vertelde die die stappen met business case, ontwikkeling...

00:39:50.258 --> 00:40:04.268

Machine Learning Engineer/Projectmanager

Ja ja ja en en en ja precies en daarbinnen hebben ook wel een stappenplan hoor, dus dus ook het hè? Dus het explainable maken van Van een voorspelling zodat zo'n operator makkelijk aan de slag kan kom je veel tegen.

00:40:04.948 --> 00:40:13.188

Machine Learning Engineer/Projectmanager

concept drift komen steeds meer tegen, dus het kijken naar of het domein waarin het model opereert, ook wel domein is waar ze bedoeld is, zoals met dat beursklimaat.

00:40:14.988 --> 00:40:18.088

Machine Learning Engineer/Projectmanager

Ja een goede business case opstellen? Ja, dus dat zijn zeker allemaal vinkjes.

00:40:19.008 --> 00:40:19.748

Anouk Wolters

Ja oké.

00:40:20.448 --> 00:40:25.418

Anouk Wolters

En worden de stappen die gemaakt worden en keuzes die gemaakt worden ook gedocumenteerd?

00:40:27.278 --> 00:40:34.198

Machine Learning Engineer/Projectmanager

Ja, We hebben, We hebben gewoon een documentatie die iets meer operationele focus heeft dan precies alle.

00:40:35.768 --> 00:40:42.068

Machine Learning Engineer/Projectmanager

Alle waaroms, maar gewoon welke data we gebruiken, in welke situaties wel en niet gebruikt. Dat is allemaal gedocumenteerd, ja.

00:40:43.188 --> 00:40:48.798

Anouk Wolters

Is dat ook standaard wat er altijd gedocumenteerd moet worden, of verschilt het per use case?

00:40:52.318 --> 00:40:59.608

Machine Learning Engineer/Projectmanager

Nee, Ik denk dat het zeker goed is om de om altijd ook open te zijn over de zwakke punten van het systeem is ontwikkeld. Ja.

00:41:01.108 --> 00:41:04.818

Anouk Wolters

Dus hebben jullie ook een standaard manier waarop documentatie.

00:41:05.618 --> 00:41:06.358

Anouk Wolters

Gedaan wordt?

00:41:06.918 --> 00:41:14.298

Machine Learning Engineer/Projectmanager

Nee, nee, dat elke klant heeft ook andere behoeften. Dat heeft ook heel erg te maken met op welk niveau je ook moet documenteren, bijvoorbeeld dus.

00:41:15.468 --> 00:41:19.878

Machine Learning Engineer/Projectmanager

Ja soms ik een heel technische documentatie overhandigen. Bij sommige organisaties juist niet.

00:41:22.418 --> 00:41:26.558

Anouk Wolters

Dus dat niveau wordt dan bepaald door wat de klant wil of hoe wordt dat bepaald?

00:41:27.848 --> 00:41:28.978

Machine Learning Engineer/Projectmanager

Ja door wat de klant?

00:41:29.758 --> 00:41:31.358

Machine Learning Engineer/Projectmanager

Dus eigenlijk door wie je klant is.

00:41:31.828 --> 00:41:32.188

Anouk Wolters

OK.

00:41:32.038 --> 00:41:47.558

Machine Learning Engineer/Projectmanager

Dus, dus zijn dat hele technische Mensen. Het zijn de Mensen die heel zwaar betrokken zijn, of zijn dat Mensen die helemaal niet betrokken zijn, zijn dat Mensen die graag documentatie willen ja of nee door klanten, zegt, die zegt ja documentatie verouderd toch? Dus dat hoeft voor ons niet zo. Dus Als je het gewoon even in een paar regels In de code kunt zetten, is dat ook prima.

00:41:49.878 --> 00:41:50.328

Machine Learning Engineer/Projectmanager

Ja.

00:41:55.538 --> 00:41:59.818

Anouk Wolters

Kijk, en je zegt documentatie verouderd, toch?

00:42:01.068 --> 00:42:06.458

Anouk Wolters

Als er wel wordt gedocumenteerd, gebeurt er dan vervolgens nog iets met deze documentatie of wordt dat geupdate?

00:42:07.668 --> 00:42:09.828

Anouk Wolters

Hoe wordt daar over de tijd mee omgaan?

00:42:10.138 --> 00:42:20.448

Machine Learning Engineer/Projectmanager

Ja er wordt geupdate dus dus bij [de bank] hebben we altijd gewoon als we een nieuwe stap toevoegd hebben hebben we wel geupdate? Maar ik vraag me af of die documentatie gebruikt wordt, denk dat dat toch nog meer In de hoofden van Mensen zit.

00:42:21.638 --> 00:42:22.278

Anouk Wolters

Ja OK.

00:42:23.548 --> 00:42:24.538

Anouk Wolters

En.

00:42:26.798 --> 00:42:27.758

Anouk Wolters

Even kijken.

00:42:30.058 --> 00:42:31.018

Anouk Wolters

Zijn.

00:42:32.058 --> 00:42:35.978

Anouk Wolters

Is model output van dit model reproduceerbaar?

00:42:42.638 --> 00:42:43.338

Machine Learning Engineer/Projectmanager

Hoe bedoel je?

00:42:43.828 --> 00:42:46.618

Anouk Wolters

Nou, als er model output is en.

00:42:47.528 --> 00:42:53.368

Anouk Wolters

Ja er wordt, dan kan dan worden gereconstrueerd hoe die model output daar komt, wordt alles gelogd?

00:42:54.688 --> 00:42:55.058

Machine Learning Engineer/Projectmanager

Ja.

00:42:55.788 --> 00:42:56.118

Anouk Wolters

Ja.

00:42:56.558 --> 00:42:56.918
Machine Learning Engineer/Projectmanager
Ja.

00:42:58.488 --> 00:42:59.218
Machine Learning Engineer/Projectmanager
Ja zeker.

00:43:01.328 --> 00:43:02.748
Anouk Wolters
En even kijken.

00:43:06.138 --> 00:43:11.248
Anouk Wolters
Wat denk je dat de grootste uitdagingen waren tijdens het ontwikkelen van deze use case?

00:43:20.718 --> 00:43:21.858
Machine Learning Engineer/Projectmanager
Ik denk ja.

00:43:25.428 --> 00:43:26.148
Machine Learning Engineer/Projectmanager
Zorgen dat.

00:43:32.238 --> 00:43:33.128
Machine Learning Engineer/Projectmanager
Ja eigenlijk.

00:43:36.548 --> 00:43:38.298
Machine Learning Engineer/Projectmanager
Dat je model het niet te goed gaat doen.

00:43:39.348 --> 00:43:47.508
Machine Learning Engineer/Projectmanager
Dus je je modelleert Mensen uiteindelijk, je modelleert Mensen hun gedrag en wanneer jij denkt dat het het juiste moment is om ze te mailen.

00:43:48.248 --> 00:44:01.098
Machine Learning Engineer/Projectmanager
Dus als jij een model krijgt wat 100% output geeft en zegt, Ik weet 100% zeker ja dat is niet zo. Er zit gewoon veel meer achter bij mensen dan jij een data kunt vangen en zeker bij de data die wij hadden die heel oppervlakkig was.

00:44:01.788 --> 00:44:08.878
Machine Learning Engineer/Projectmanager
Dus Het was een soort Als ik zag dat het model te goed ging doen, dan wist je eigenlijk al zeker dat er iets verkeerd aan de hand was.

00:44:09.628 --> 00:44:26.808

Machine Learning Engineer/Projectmanager

Dus Het was heel ja, een beetje ambigu om te vinden waar dan de grens zat tussen. Oké, Ik denk nu dat die iets heeft gevonden waar we op af kunnen gaan, maar ook goed generaliseert. En niet zo specifiek is dat die overfit is. Dat is denk ik, Dat is hier heel belangrijk.

00:44:27.868 --> 00:44:28.318

Anouk Wolters

Oké.

00:44:30.908 --> 00:44:31.338

Anouk Wolters

OK.

00:44:32.288 --> 00:44:32.638

Anouk Wolters

even Kijken

00:44:33.698 --> 00:44:39.908

Anouk Wolters

Denk dat ik verder op zich wel door de vragen heen bent, denk jij dat we nog iets hebben gemist in het interview?

00:44:41.858 --> 00:44:42.768

Anouk Wolters

Zou je nog iets willen doen?

00:44:53.388 --> 00:44:57.748

Machine Learning Engineer/Projectmanager

Nou ja, kijk, Ik denk maar waar [de bank]

00:44:59.418 --> 00:45:00.618

Machine Learning Engineer/Projectmanager

mee worstelt, met name ook nu.

00:45:01.368 --> 00:45:01.948

Machine Learning Engineer/Projectmanager

Is wat

00:45:02.858 --> 00:45:06.528

Machine Learning Engineer/Projectmanager

Wat doe je met zo'n project Als je, hoe zorg je ervoor dat je echt

00:45:10.688 --> 00:45:21.418

Machine Learning Engineer/Projectmanager

kunt internaliseren, dus hoe zorg je ervoor dat je onafhankelijk van [external ML

engineering/consulting firm] zo'n model kunt blijven gebruiken? En dan ook stel je voor, wij gaan failliet en wij verdwijnen van deze wereld af.

00:45:22.028 --> 00:45:42.848

Machine Learning Engineer/Projectmanager

Kun je dat model dan nog blijven gebruiken en kun je ook Als het fout gaat of als dingen die je niet begrijpt dan daar mee omgaan? Maar ik denk dat [de bank] daar nu nog niet voldoende vertrouwen heeft, dus ze zouden graag het model gewoon graag helemaal intern houden en met ons andere dingen gaan doen, Maar dat lukt ze gewoon nog niet helemaal Omdat ze zelf ook nog niet de juiste kennis in huis hebben. En Dat is ook deels Natuurlijk Waarom je consultants aanneemt.

00:45:43.698 --> 00:45:53.458

Machine Learning Engineer/Projectmanager

Maar dat is volgens mij altijd een spanningsveld dat je bij bedrijven ziet, en sommige bedrijven zijn er helemaal oké mee om gewoon voor de rest van je leven. Service level agreement te hebben met [external ML engineering/consulting firm] en andere bedrijven die vinden dat lastiger.

00:45:54.168 --> 00:45:54.568

Anouk Wolters

Ja.

00:45:56.818 --> 00:45:57.328

Anouk Wolters

Oké.

00:45:58.028 --> 00:45:58.838

Anouk Wolters

Goede toevoeging.