Literatuurstudie:

# Inleiding

Artificiële intelligentie is in elk vakgebied een essentieel onderdeel om aan de top van je sector te blijven in het jaar 2024. Daarom is ook het bedrijf Itineris zich in dit onderwerp aan het verdiepen. Itineris is een bedrijf dat zich volledig inzet op het vergemakkelijken van de business processen aan de hand van digitale tools voor bedrijven die handelen met basisvoorzieningen zoals water en elektriciteit.

Itineris zet de nadruk op efficiëntie aan de hand van geautomatiseerde processen, een versimpelde UX (User Experience) en de mogelijkheid aan klanten om zichzelf te helpen. Dit resulteert voor de klanten dan in een gegarandeerde lagere service kost.

Itineris zorgt ook voor betere inzichten voor de klanten aan de hand van omni-channel features die je toelaten van proactief de tevredenheid van je klant in de goede richting te sturen. Je kan business processen vergemakkelijken aan de hand van een flexibele, snelle en super aanpasbare third-party integration.

De software die Itineris aanbied maakt gebruik van het MS-Dynamics 365 ERP-pakket. Microsoft Dynamics 365 is een portfolio met intelligente zakelijke applicaties die superieure operationele efficiëntie koppelt aan baanbrekende klantervaringen – waardoor jij meer agile te werk kunt gaan en de complexiteit kunt verminderen zonder stijgende kosten. Wat Itineris onderscheid van de competitie is dat ze bovenop dit pakket van Microsoft nog een eigen aanvullende software laag hebben gebouwd.

Itineris is heel veel bezig met customer service om de klanten zo goed mogelijk verder te helpen en ze doen dit op verschillende manier zoals bv: telefonische gesprekken, e-mails, chats en dergelijke. Ze doen dit al op een hoog niveau maar ze blijven naar verbetering zoeken. Dit is dan ook de reden dat ze nu naar manieren kijken om artificiële intelligentie te gebruiken om deze processen te vergemakkelijken en efficiënter te laten verlopen. Zo zouden ze bijvoorbeeld een chatbot willen die vanzelf klanten kan helpen aan de hand van bestaande documentatie zonder dat de medewerkers van Itineris er zelf zouden tussen moeten komen. Maar ze kijken niet enkel naar de client-side, ze willen ook een copilot die de medewerkers kan helpen met het uitvoeren van simpele business opdrachten. Deze copilot zou dan moeten kunnen helpen met het verwerken van tickets of snel medewerkers moeten kunnen helpen als ze iets zouden moeten opzoeken.

Momenteel zijn ze dit al aan het bekijken hoe ze dit kunnen implementeren op andere vlakken binnen het bedrijf maar nog niet bij de customer service. Ze werken hiervoor met de copilot die Microsoft aanbiedt, Microsoft copilot 365, maar hiermee komen ze nog bij enkele dingen in de knoop. Sommige dingen kan het standaard pakket van de Microsoft copilot al maar de functionaliteiten die gebruik zouden maken van de extra laag die bovenop het MS Dynamics pakket is gebouwd werken nog niet. Dit zouden we dus verder onderzoeken hoe we deze copilot kunnen trainen en implementeren zodat deze op de correcte manier de relevante informatie zou kunnen voorleggen.

Op basis hiervan heb ik de volgende onderzoeksvraag geformuleerd: *onderzoeksvraag: Hoe kunnen we de efficiënte ontwikkeling van een co-pilot bevorderen door het identificeren van verschillende valkuilen en trainingsmethoden met specifieke aandacht voor de vereisten van de onderneming?*

Om op deze onderzoeksvraag een kwalitatief antwoord te kunnen geven moeten we uiteraard op een gestructureerde manier onderzoek doen in de relevante onderwerpen. De beste manier om dit te voltooien is van op voorhand de nodige doelstellingen op te stellen in overleg met het bedrijf en een de hand daarvan aan de slag te gaan.

Ik heb voor de casus van Itineris de volgende doelstellingen opgesteld:

* Identificeren van de belangrijkste vereisten voor de copilot
  + Wat moet de copilot kunnen?
  + Wat zijn de criteria om te beslissen of een handeling succesvol is?
* Identificeren van de verschillende valkuilen die we kunnen tegenkomen bij het ontwikkelen van de copilot.
  + Wat zijn de gekende valkuilen bij het ontwikkelen van een copilot?
  + Hoe kunnen we deze valkuilen het beste omzeilen?
* Trainen van de copilot.
  + Hoe train je een copilot?
  + Waar zit de informatie die de copilot moet verwerken?
  + Hoe gaat de copilot trainen met deze informatie?
* Ontwikkelen van een prototype van een copilot die voldoet aan de criteria van punt 1.

# Stand van zaken

Het onderzoeksdomein van artificiële intelligentie is al een stuk langer aanwezig dan wat de meeste mensen denken. Volgens een artikel van The Council of Europe is het nog steeds een relatief “jong” onderzoeksdomein van 60 jaar oud terwijl de meeste mensen er maar in contact mee zijn gekomen de laatste 10 jaren. Het onderzoek naar artificiële intelligentie, wat een set is van wetenschappen, theorieën en technieken (inclusief wiskundige logica, statistiek, waarschijnlijkheden, computationele neurobiologie, informatica), richt zich op het imiteren van cognitieve vaardigheden van de mens. Geïnstantieerd in de tweede wereldoorlog, is de ontwikkeling sterk gelinkt aan die van computers en heeft geleid naar de mogelijk om computers steeds ingewikkeldere taken te laten uitvoeren, wat we hiervoor enkel aan mensen konden overlaten.

Desondanks het nog steeds ver van menselijke intelligentie staat in de strikte vorm is artificiële intelligentie toch al sterk ontwikkeld. Het ultieme doel is van een “sterke” artificiële intelligentie te ontwikkelen die sterk verschillende gespecialiseerde problemen op een autonome manier kan oplossen. Momenteel zitten we hier nog helemaal niet aan aangezien de artificiële intelligentie die nu op de markt te vinden is enkel nog maar in staat is van efficiënt te helpen in de specifieke gebieden waarbinnen ze getraind zijn. Sinds 2010 heeft de discipline een sterke verbetering ondervonden door de grote verbetering in kracht van computers en toegang tot grote hoeveelheden data.

## Machine learning

Machine learning is een vorm van artificiële intelligentie dat gericht is op het bouwen van systemen die van verwerkte data kunnen leren of data gebruiken om prestaties te verbeteren. De termen artificiële intelligentie en machine learning worden vaak door elkaar gebruikt terwijl dit twee verschillende dingen zijn. Artificiële intelligentie is de overkoepelende term waaronder machine learning valt.

Er bestaan verschillende soorten machine learning algoritmes waarvan de meest prominente supervised machine learning , unsupervised machine learning, semi-supervised machine learning en reinforcement machine learning zijn.

### Supervised machine learning algoritmes

Supervise machine learning algoritmes zijn afhankelijk van correct gelabelde data. Het algoritme wordt de data gegeven tezamen met de verwachte uitkomst van de data. Het model leert dan ui de relaties van de input en output data om een model te bouwen. Het model mapt de input data naar de gewenste output en wordt getraind tot het model het gewenste niveau van accuraatheid bekomt. Dit wordt altijd getraind onder de supervisie van een developer.

Supervised machine learning wordt gebruikt voor:

* Het classificeren van nieuwe data in bestaande groepen en categorieën
* Voorspellen van ontwikkelende en toekomstige trends gebaseerd op voorspellende modellen
* Bedrijven helpen in campaign en acquisitie projecten door het voorspellen van veranderingen van variabelen

### Unsupervised machine learning algoritmes

Unsupervised machine learning algoritmes worden niet gecontroleerd door een developer en worden getraind op datasets zonder labels. Unsupervised machine learning algoritmes worden gebruikt voor het herkennen van patronen, trends of groepen van datasets waarvan de elementen onbekend zijn. Deze soort machine learning kan gebruikt worden om relaties te leggen tussen verschillende data punten en het segmenteren van gelijkaardige data.

Unsupervised machine learning algoritmes worden gebruikt voor:

* Het segmenteren van data voor gepersonaliseerde marketing campaigns
* In het licht zetten van onbekend trends bij de klanten
* Gepersonaliseerde content presenteren gebaseerd op interesses van online muziek en streaming services.

### Semi-supervised machine learning algoritmes

Zoals de naam suggereert is dit een mengeling van supervised en unsupervised machine learning algoritmes. Het wordt gebruikt voor datasets waarbij slechts een deel van de data correct is gelabeld. Semi-supervised machine learning algoritmes worden getraind op de subset van correct galebelde data. Het model gebruikt dan deze training om de rest van de data te labelen.

Semi-supervised machine learning algoritmes worden gebruikt om:

* Groeperen van grote tekstdocumenten zoals gescande boeken of documenten.
* Categoriseren van data dat tijdsintensief zou zijn voor een menselijke specialist.

### Reinforcement machine learning

Reinforcement machine learning laat een systeem toe van te leren en verbeteren van een functie door trial and error. Het model start vanaf nul en zal de beste oplossing proberen vinden in een specifieke omgeving door te leren van de acties die het in vorige fases heeft genomen. Dit model zit dus eigenlijk in een feedback loop waarbij elke actie ofwel wordt beloond ofwel afgestraft. Deze iteraties worden complexer als de complexiteit van de omgeving en acties die het model kan ondernemen toenemen. Reinforcement machine learning algoritmes worden vooral gebruikt als een actie te complex is voor een statische functie die een developer heeft geschreven.

Reinforcement machine learning wordt gebruikt bij:

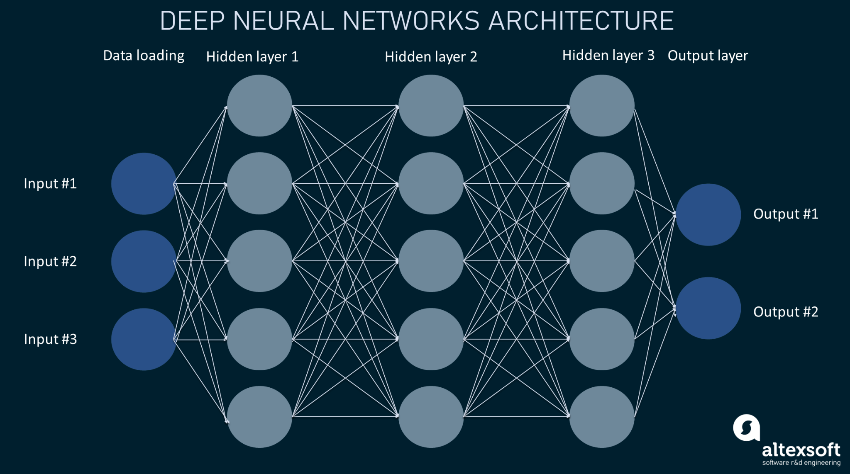
* Ontwikkeling van artificiële intelligentie
* Systemen in zelfrijdende auto’s
* Een model een spel te leren zoals schaken tegen een menselijke tegenstander

## Deep learning

Deep learning is een subset van machine learning, waarbij kunstmatige neurale netwerken (algoritmen die zijn ontworpen om te werken zoals het menselijke brein) leren van grote hoeveelheden gegevens. Er wordt van deep learning gesproken als een neuraal netwerk bestaat uit meer dan 3 lagen.

## Neurale netwerken

Een neuraal netwerk is een machine learning programma, of model, dat beslissingen maakt op een manier die vergelijkbaar is met het menselijke brein. Door gebruik te maken van processen na te bootsen die de biologische neuronen gebruiken om samen te werken voor dingen te identificeren en de opties af te wegen komen deze netwerken aan conclusies.

Elk neuraal netwerk bestaat uit lagen van nodes, of artificiële neuronen – een input laag, een of meerdere verborgen lagen en een uitput laag. Elke node staat in connectie met andere nodes en heeft zijn eigen “gewicht” en “drempelwaarde”. Als de output van een node boven de gespecifieerde drempelwaarde is, dan is de node geactiveerd, stuurt de node data door naar de volgende laag van het netwerk.

Eens de input laag is vastgesteld wordt er voor elke node een gewicht vastgesteld. Dit gewicht bepaalt het belang van de node, waarbij de nodes met een groter belang meer invloed hebben op het eindresultaat in vergelijking met de nodes met een kleiner gewicht. Voor elke node wordt de input door een activatie-functie gestuurd om de output waarde te berekenen.

Indien deze output waarde groter is dan de drempelwaarde wordt data doorgestuurd naar de volgende laag in het neurale netwerk (bijvoorbeeld van input node #1 naar alle nodes in de verborgen laag 1 die in verbinding staan met input #1). Zo wordt de output van de ene node de input van de andere.

Terwijl we het model trainen zullen we de accuraatheid van het model willen evalueren en dit doen we aan de hand van een kost functie. Dit wordt ook vaak de Mean Squared Error (MSE) genoemd. Het doel is van deze kost functie zo klein mogelijk te krijgen om de correctheid van elke observatie te verzekeren. Het model past het gewicht en de drempelwaardes aan door verschillende iteraties aan de hand van de kost functie en reinforcement learning om zo tot een punt te komen dat we het lokale minimum kunnen noemen.

## Large Language Models

Een Large Language Model (LLM) is een deep learning algoritme dat een reeks natuurlijke taalverwerkingstaken kan uitvoeren. De modellen maken gebruik van transformator modellen en zijn getraind op grote datasets. Zo kunnen ze teksten of andere content herkennen, vertalen, voorspellen of genereren. Naast het aanleren van menselijke talen aan toepassingen van artificiële intelligentie, kunnen large language models ook getraind worden om verschillende taken uit te voeren, zoals het begrijpen van eiwitstructuren, het schrijven van softwarecode en nog veel meer.

### Wat is een transformatormodel?

Een transformatormodel is de meest voorkomende architectuur van een groot taalmodel. Deze bestaat uit twee kerncomponenten een encoder en een decoder. Een transformatormodel verwerkt gegevens door invoer in tokens om te zetten en vervolgens gelijktijdige wiskundige vergelijkingen uit te voeren om de relaties tussen de tokens te ontdekken. De computer kan dus de patronen zien die een mens zou vaststellen als dezelfde vraag aan hem zou worden voorgelegd.

Transformer-modellen vertrouwen op zelfaandachtsmechanismen, waardoor ze sneller kunnen leren dan traditionele modellen, zoals modellen voor het lange kortetermijngeheugen (LSTM). Een zelfaandachtig transformatormodel kan naar verschillende delen van een reeks, of de gehele context van een zin, kijken om voorspellingen te genereren.

#### Zelfaandachtsmechanisme

De kern van het transformer-model ligt in het zelfaandachtsmechanisme. Hierdoor kan het model het belang van een token in een reeks afwegen op basis van de relevantie ervan voor andere tokens. Hiermee kan het model aandachtscores berekenen en op basis hiervan zal het model hogere gewichten toekennen aan tokens die semantisch gerelateerd zijn (een gelijkaardige/ gelinkte betekenis hebben).

#### Encoder

De encoder is verantwoordelijk voor het verwerken van de invoerreeks en de correcte representaties eruit te halen. Het bestaat uit een stapel identieke lagen, elke laag bestaat uit een zelfaandachtsmechanisme gevolgd door een feed-forward neuraal netwerk. Het zelfaandachtsmechanisme laat het model concentreren op verschillende delen van de invoerreeks en het feed-forward neuraal netwerk verwerkt de informatie lokaal. Door deze samenwerking worden niet-lineaire transformaties mogelijk gemaakt.

#### Decoder

De decoder accepteert de gecodeerde representaties van de encoder en genereert een uitvoerreeks. Het bevat ook zelfaandachtsmechanismen om afhankelijkheden binnen de uitvoerreeks vast te leggen en aandacht te besteden aan de relevante delen van de invoerreeks. Hierdoor kan het model contextueel bewuste uitput genereren voor taken zoals automatische vertaling of samenvatting van teksten.

## Copilots

Een AI copilot is een interface die gebruik maakt van Large Language Models (LLMs) om gebruikers te ondersteunen in verschillende handelingen en beslissende processen. Door gebruik te maken van de LLMs kunnen AI copilots grote hoeveelheden data begrijpen, analyseren en behandelen.

Ai copilots spelen een grote rol in het verbeteren van productiviteit en efficiëntie van handelingen door:

* Het bieden van context gevoelige assistentie:
  + Ai copilots kunnen proactief reageren op de handelingen van de gebruiker aan de hand van een bepaalde context.
  + Het verzekerd relevantie en snelle reactietijden voor de nodige processen.
* Automatiseren van alledaagse taken:
  + Door repetitieve en tijds consumerende taken in handen te nemen laat de copilot toe van de gebruikers hun tijd te steken in strategische en creatieve taken dat de productiviteit doet stijgen.
* Analyseren van data:
  + AI copilots kunnen snel grote hoeveelheden data verwerken, patronen en trends herkennen, en correcte inzichten op deze data weergeven voor beslissingen te maken.
* Verbeteren van communicatie:
  + AI-copiloten faciliteren effectieve interacties met verschillende belanghebbenden, waaronder werknemers, klanten en leveranciers, waardoor communicatieprocessen worden gestroomlijnd en vertragingen of misverstanden worden verminderd.
* Ongelijksoortige systemen te laten samenwerken:
  + Copilots kunnen de “tussenpersoon” zijn voor verschillende platformen, tools of software te laten samenwerken onder een geheel met de zekerheid van het behoud van data, toegang en compatibiliteit over het geheel.

Samengevat staat een een copilot in om handelingen te versimpelen en geldige begeleiding te bieden, met als doel de user experience te verbeteren en de business goals effectief en efficiënt te voltooien.

## Microsoft copilot

De copilot die Itineris zou willen implementeren is de copilot die Microsoft in 2023 heeft uitgebracht namelijk de Microsoft copilot. Het is gebaseerd op de Large Language Models en is daardoor in staat van verschillende dingen uit te voeren zoals bronnen citeren, liedjes schrijven, etc.. .

Het Maakt gebruik van het Microsoft Prometheus model, gebouwd bovenop GPT-4 van OpenAI. Dit is een fundamenteel Large Language Model dat is gefinetuned aan de hand van supervised en reinforcement learning technieken. De chatbot is dan ook in staat van te communiceren in verschillende talen en dialecten zoals ChatGPT van OpenAI.