

NVB Risicogebaseerde Standaard

Technische model- documentatie

Inleiding

Banken maken steeds vaker gebruik van advanced analytics, kunstmatige intelligentie (AI) en machine learning modellen bij de naleving van Wwft-verplichtingen. Deze modellen kunnen op grote schaal binnen het financial crime framework worden toegepast, zoals bij het onboarden van klanten, voortdurende controle (Ongoing Due Diligence; hierna: ODD) en transactiemonitoring. Innovaties op dit gebied maken het voor banken mogelijk om hun detectieprocessen voor risico's op witwassen en terrorismefinanciering (hierna: ML/TF) effectiever en efficiënter in te richten. Ook bieden ze mogelijkheden om de uitvoering van anti-witwassen en terrorismefinanciering (hierna: AML/CFT) en risk management processen effectief en adequaat te verbeteren.

Aangezien incidenten als gevolg van het gebruik van modellen serieuze gevolgen voor de reputatie van banken en de sector kunnen hebben, gelden voor banken een hoge integriteitsstandaard. Banken zijn verantwoordelijk voor hun gebruik van advanced analytics, AI en machine learning, omdat toepassingen niet altijd functioneren zoals bedoeld en kunnen leiden tot nadelen, schade of verlies voor banken of hun klanten. Modeltoepassingen moeten daarom betrouwbaar zijn, zich voorspelbaar gedragen en werken binnen de geldende wet- en regelgeving.

Deze NVB Standaard beschrijft uitgangspunten voor banken in Nederland om aan te tonen dat passende processen zijn gevolgd en om de betrouwbaarheid van de modellen te waarborgen. Het opstellen van technische modeldocumentatie draagt bij aan verantwoording en transparantie, wat betekent dat banken uitleggen hoe en waarom ze advanced analytics, AI en machine learning modellen gebruiken in hun bedrijfsprocessen en hoe deze toepassingen functioneren.

De ontwikkeling en het gebruik van advanced analytics, AI en machine learning modellen verschilt per bank, gebaseerd op hun eigen behoefte en business model. De ontwikkeling en toepassing van dergelijke modellen is ook afhankelijk van de klantportefeuille, mate van technische volwassenheid en de beschikbare capaciteit van de bank. De in deze NVB Standaard beschreven maatregelen geven een algemeen overzicht van de vereiste informatie en documentatie voor verschillende typen modellen wanneer deze worden ontwikkeld en ingezet in het AML/CFT framework. Het beschrijft criteria voor modelkeuzes, waaronder uitlegbaarheid, redelijkheid en betrouwbaarheid van de gekozen modellen.

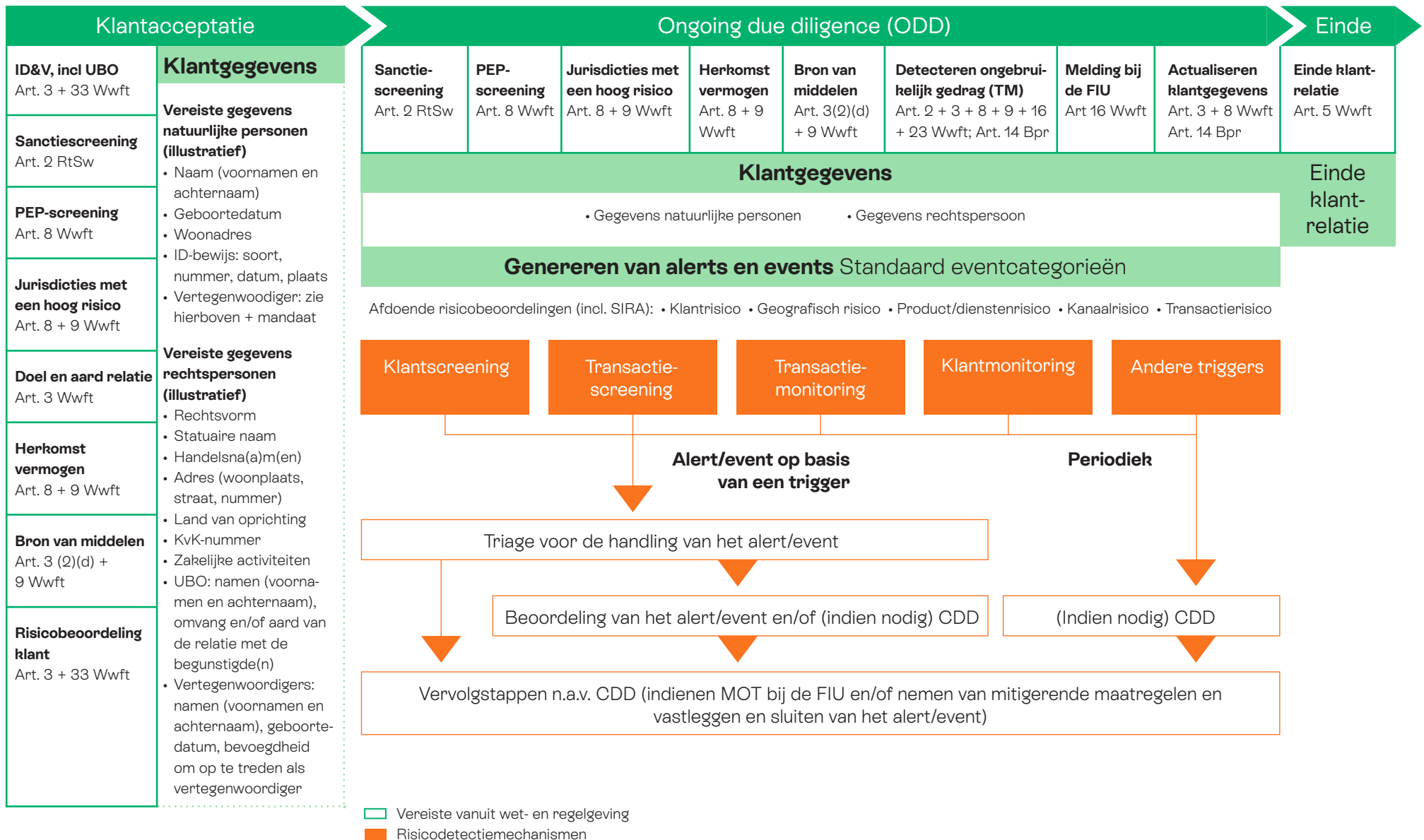
Deze NVB Standaard moet worden gelezen in samenhang met de NVB Standaard 'Modellen voor het genereren van alerts en events', die een overzicht geeft hoe modellen kunnen worden gebruikt in een risicogebaseerde benadering om ML/TF risico's te detecteren.

De positionering binnen het Financial Crime Framework

Modellen kunnen door banken worden toegepast in verschillende stadia van het Financial Crime Framework, zoals bij onboarding, voor ODD processen en transactiemonitoring. In dit risk management framework gebruiken modellen doorgaans klant- en transactiegegevens als input en genereren uitkomsten die de basis vormen voor de risico-beoordeling van de klant en het genereren van alerts en events. De dekking van ML/TF risico's, zoals beschreven in de risk assessment van banken (zoals de Systematic Integrity Risk Analysis), is een belangrijke overweging.

Financial Crime Framework

Risicogebaseerd



NVB Risicogebaseerde Standaard

1 NVB Standaard

Deze NVB Standaard beschrijft welke soorten informatie en documentatie moeten worden uitgewerkt bij de ontwikkeling en het gebruik van advanced analytics, AI en machine learning modellen in de context van het AML/CFT framework. Met deze technische modeldocumentatie kunnen banken aantonen dat er passende processen zijn gevolgd en kan de betrouwbaarheid van advanced analytics, AI en machine learning modellen worden gewaarborgd. De documentatie biedt ook transparantie en kan in de toekomst als referentie dienen bij het updaten en valideren van modellen.

Analytics modellen zijn er in verschillende soorten. Het is belangrijk om voldoende capaciteit beschikbaar te hebben voor advanced analytics, AI en machine learning methoden en technieken om te zorgen dat de ontwikkelde en toegepaste advanced analytics, AI en machine learning modellen adequaat zijn en ‘well-understood’. Verantwoording en oversight zijn belangrijke elementen bij advanced analytics, AI en machine learning. Het verantwoord gebruik van deze modellen moet worden ondersteund door het governance en model risk management framework van banken waarbij de technische modeldocumentatie kan dienen om relevante technische, compliance en ethische risico's te kunnen identificeren en mitigeren.

De technische modeldocumentatie in deze Standaard, biedt een framework dat banken ondersteunt bij het aantonen hoe hun toepassing van advanced analytics, AI en machine learning modellen wordt ontwikkeld en geëvalueerd. De technische modeldocumentatie kan dienen om op een betrouwbare, kwalitatieve, objectieve en meetbare wijze over de modelperformance te rapporteren. Dit maakt een betrouwbaar, transparant en verantwoord gebruik van deze modellen mogelijk. Bovendien kan hiermee de ‘fairness’, uitlegbaarheid en interpreteerbaarheid van het model worden aangetoond.

De technische modeldocumentatie voorziet banken van een framework dat ervoor zorgt dat advanced analytics, AI en machine learning modellen worden gebruikt waarvoor ze bedoeld zijn.

De goede werking van applicaties gebaseerd op advanced analytics, AI en machine learning, moet regelmatig worden beoordeeld. Vooral voor scenario's waarbij het model zich onverwacht gedraagt, moet er een ‘incident response plan’ zijn met stappen voor snelle detectie, respons en herstel. Issues met integriteit van data en bias, zowel in de ontwikkeling als in de productie, moeten worden geëvalueerd en op een structurele manier worden gedocumenteerd voor toekomstig gebruik. Op dezelfde manier kunnen feedback

loops worden toegepast om ervoor te zorgen dat veranderende patronen in input of output in een vroeg stadium worden gesignaleerd, zodat modellen dienovereenkomstig kunnen worden aangepast en effectief blijven.

De technische modeldocumentatie past binnen het algehele governance framework dat ziet op de levenscyclus van advanced analytics, AI en machine learning modellen. Dit omvat de rollen en verantwoordelijkheden van verschillende interne en externe stakeholders die betrokken zijn bij de ontwikkeling, inzet en het onderhoud van de modellen. De documentatie maakt ook verantwoording en communicatie met verschillende stakeholders mogelijk over de mogelijkheden en beperkingen van de modellen. Daarnaast kan het als leidraad dienen voor het trainen van medewerkers om de output van de modellen te begrijpen en op de juiste wijze te interpreteren.

De technische modeldocumentatie moet het volgende vaststellen en definiëren:

- **model governance:** alle governance rollen en processen;
- **conceptueel modelontwerp:** alle belangrijke beslispunten gemaakt met de business (1e-lijn) die van invloed zijn geweest op het ontwerp;
- **technisch modelontwerp en data:** beschrijving van brondata, datakwaliteit, data ‘processing’- en ‘pre-processing’-stappen;
- **modelontwikkeling:** gebruikte methodiek, validatie van resultaten en gebruikte stabiliteits-tests;
- **uitlegbaarheid:** het vermogen om uit te leggen wat er in een model gebeurt van input naar output;
- **risico-overwegingen:** overzicht van alle risico’s en van beperkingen van gebruikte model(len) en de daarbij gebruikte technologieën;
- **modelonderhoud in productie:** plan voor het monitoren en door ontwikkelen van de performance van het model in productie;
- **implementatie van het model:** hoe het model in productie wordt genomen.

Zie de bijlage voor een meer gedetailleerde omschrijving van de technische modeldocumentatie.

2 Impact

Advanced analytics, AI en machine learning modellen kunnen een aanzienlijke bijdrage leveren aan een efficiënt AML/CFT framework omdat hiermee de informatiepositie van banken over klanten en transacties kan worden verbeterd en ML/TF-risico’s accuraat kunnen worden gedetecteerd. Deze systemen kunnen onnodige lasten voor klanten verminderen en risicorelevante AML/CFT-maatregelen verbeteren.

Advanced analytics, AI en machine learning en innovatie brengen daarentegen ook additionele verantwoordelijkheden en verantwoording voor banken met zich mee. Er is voldoende menselijke tussenkomst en oversight nodig voordat beslissingen worden genomen met betrekking tot een klant. De impact op de samenleving en klanten kan groot zijn als advanced analytics, AI en machine learning niet verantwoord worden toegepast. De veiligheid en fundamentele rechten van mensen en organisaties moeten worden gewaarborgd door biases te voorkomen en risico’s op inbreuk van het recht op privacy te vermijden.

De technische modeldocumentatie voor deze systemen zoals uiteengezet in deze NVB Standaard zal bijdragen aan een consistente aanpak voor het ontwikkelen, evalueren en gebruiken van advanced analytics, AL en machine learning modellen.

3 Use cases

De onderstaande use case is een voorbeeld om een praktische toepassing van deze NVB Baseline te illustreren en niet bedoeld om volledig te zijn.

Combinatie van een rule-based model en machine-learning

Voorbeeld

Een bank wil haar rule-based model voor transactiemonitoring verbeteren naar een risicogebaseerd model dat rule-based systeem combineert met machine learning.

NVB Standaard

- Gedurende de risk assessment heeft een bank 20 red flags geïdentificeerd die zouden kunnen duiden op ML risico’s. Deze risico’s zijn tot nu toe benaderd vanuit een rule-based transactiemonitoring set-up, wat resulteerde in onnauwkeurige uitkomsten (d.w.z. een groot aantal false positives). De bank besluit de mogelijkheid te onderzoeken om het probleem te tackelen met een machine learning model.
- De bank stelt een groep samen van business, compliance en modelling experts die overeenstemming bereiken over het doel en ontwerp van het model.
- Als onderdeel van het implementatietesten vergelijkt de bank de eerste uitkomsten van het model met de uitkomsten van de bestaande business rules over de afgelopen 2 jaar. Het model kan 93 van de 96 gevallen identificeren

die eerder door het rule-based systeem zijn gesignaleerd en die zijn geëscaleerd en gerapporteerd. Slechts 3 van de verdachte activiteiten (Suspicious Activity Reports; hierna SAR) worden gemist. Maar het model vindt ook nog eens 16 SAR-waardige cases.

- De compliance expert van de groep voert een analyse uit van de 3 gemiste SARs en concludeert dat 2 van de 3 SARs bijvangst waren, enkel gebaseerd op niet-transactie-gerelateerde informatie. De derde gemiste SAR is daarentegen een complexe case met duidelijke ML-risico's die aandacht vereisen. De groep besluit om de SAR samen te analyseren, maar is niet in staat om een gerichte manier te vinden om de SAR met het nieuwe model te detecteren.
- De groep legt hun bevindingen vast en stelt voor dat de voordelen van het model opwegen tegen het missen van één enkele SAR.
- Na goedkeuring van Compliance wordt het model parallel aan de regels in het productie-proces geïmplementeerd om zo technische en performance stabiliteit te garanderen.
- Na het model een maand parallel aan de huidige set-up te hebben gedraaid, is het team er zeker van dat het model zonder problemen werkt en klaar is voor implementatie.
- Het voorstel wordt ingediend bij het Compliance comité, waarbij de voor- en nadelen worden uiteengezet en het definitief wordt goedgekeurd.
- Het model wordt geïmplementeerd en parallel daaraan start de monitoring van het model.

Het regelgevend kader

De regelgevende context die van toepassing is op dit onderwerp is te vinden in de relevante delen van de toepasselijke wetten, regelgeving en in richtlijnen van verschillende autoriteiten. De Wwft vereist dat banken de integriteitsrisico's van klanten beoordelen bij onboarding en doorlopend tijdens de zakelijke relatie en ODD uitvoeren, waaronder onderzoek naar transacties die gedurende de zakelijke relatie worden verricht.

Er zijn geen wettelijke vereisten die voorschrijven of en hoe banken modellen moeten gebruiken bij het implementeren van de Wwft-verplichtingen. DNB en de Wolfsberg Group geven evenwel algemene richtlijnen voor het gebruik van advanced analytics, AI en machine learning modellen bij banken. Bovendien zal de EU Artificial Intelligence Act (AI Act) van invloed zijn op de eisen voor banken om verantwoord gebruik te maken van modellen en dus ook op deze NVB Standaard.

Het gebruik van advanced analytics, AI en machine learning modellen kan vraagstukken oproepen die buiten de scope AML/CFT vallen. Meer specifiek moeten banken voldoen aan de raakvalken met de General Data Protection Regulation (GDPR) en de AI Act.

Hieronder een overzicht van relevante wet-, regelgeving en richtlijnen.

- **General Principles for the use of AI in the Financial Sector, DNB, 2019**
“AI applications in the financial sector should be reliable and accurate, behave predictably, and operate within the boundaries of applicable rules and regulations. Firms should also be accountable for their use of AI, as AI applications may not always function as intended and can result in damages for the firm itself, its clients and/or other relevant stakeholders. it is vital for society’s trust in the financial sector that financial firms’ AI applications – individually or collectively – do not inadvertently disadvantage certain groups of clients.”
- **Wolfsberg Principles for using AI and Machine Learning in Financial Crime Compliance, The Wolfsberg Group, 2022**
“By leveraging the advances in data science underpinning AI/ML, FIs can holistically analyse the client and transactional data created by their products and services more effectively and efficiently to detect, investigate, and manage the risk of financial crime, and satisfy regulatory requirements.”
“Design of AI/ML systems should be driven by a clear definition of the intended outcomes and ensure that results can be adequately explained or proven given the data inputs.”

- **EBA Machine Learning for IRB Models, Follow-up Report from the Consultation on the Discussion Paper on Machine Learning for IRB Models, EBA/REP/2023/28, August 2023**
“Another challenge for financial institutions relates to the explainability and interpretability of the results of ML models. A higher level of complexity may lead to better model performance but at the cost of lower explainability and comprehension of the model’s functioning.”
In addition, to ensuring that the model is correctly interpreted and understood, institutions are recommended to (...) provide a summary document in which the model is explained in an easy manner based on the outcomes of the analyses described in point a. The document is recommended to describe:
 - i. The key drivers of the model.
 - ii. The main relationships between the risk drivers and the model predictions. The addressees of the document are all the relevant stakeholders, including the staff which uses the model for internal purposes.”
- **Charter of fundamental rights of the European Union, 2012**
“The peoples of Europe, in creating an ever closer union among them, are resolved to share a peaceful future based on common values. Conscious of its spiritual and moral heritage, the Union is founded on the indivisible, universal

values of human dignity, freedom, equality and solidarity; it is based on the principles of democracy and the rule of law. It places the individual at the heart of its activities, by establishing the citizenship of the Union and by creating an area of freedom, security and justice.”

- **AI Act (final text to be published) – Seven principles**

“The seven non-binding ethical principles for AI should help ensure that AI is trustworthy and ethically sound. The seven principles include: human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being and accountability.”

- **AI Act (final text to be published) Article 17**

“1. Providers of high-risk AI systems shall put a quality management system in place that ensures compliance with this Regulation. That system shall be documented in a systematic and orderly manner in the form of written policies, procedures and instructions, and shall include at least the following aspects:

- a) a strategy for regulatory compliance, including compliance with conformity assessment procedures and procedures for the management of modifications to the high-risk AI system;
- b) techniques, procedures and systematic actions to be used for the design, design control and design verification of the high-risk AI system;

- c) techniques, procedures and systematic actions to be used for the development, quality control and quality assurance of the high-risk AI system;
- d) examination, test and validation procedures to be carried out before, during and after the development of the high-risk AI system, and the frequency with which they have to be carried out;
- e) technical specifications, including standards, to be applied and, where the relevant harmonised standards are not applied in full, or do not cover all of the relevant requirements set out in Chapter II of this Title, the means to be used to ensure that the high-risk AI system complies with those requirements;
- f) systems and procedures for data management, including data acquisition, data collection, data analysis, data labelling, data storage, data filtration, data mining, data aggregation, data retention and any other operation regarding the data that is performed before and for the purposes of the placing on the market or putting into service of high-risk AI systems;
- g) the risk management system referred to in Article 9;
- h) the setting-up, implementation and maintenance of a post-market monitoring system, in accordance with Article 61;
- i) procedures related to the reporting of a serious incident in accordance with Article 62;

- j) the handling of communication with national competent authorities, other relevant authorities, including those providing or supporting the access to data, notified bodies, other operators, customers or other interested parties;
 - k) systems and procedures for record keeping of all relevant documentation and information;
 - l) resource management, including security of supply related measures;
 - m) an accountability framework setting out the responsibilities of the management and other staff with regard to all aspects listed in this paragraph.
2. The implementation of aspects referred to in paragraph 1 shall be proportionate to the size of the provider’s organisation. Providers shall in any event respect the degree of rigour and the level of protection required to ensure compliance of their AI systems with this Regulation.
- 2a. For providers of high-risk AI systems that are subject to obligations regarding quality management systems or their equivalent function under relevant sectorial Union law, the aspects described in paragraph 1 may be part of the quality management systems pursuant to that law.
3. For providers that are financial institutions subject to requirements regarding their internal governance, arrangements or processes under Union financial services legislation, the obligation to put in place a quality management system with the exception of paragraph 1, points (g), (h) and (i) shall be deemed to be fulfilled by

complying with the rules on internal governance arrangements or processes pursuant to the relevant Union financial services legislation. In that context, any harmonised standards referred to in Article 40 of this Regulation shall be taken into account.

Afstemming tussen de ‘DNB Good Practices’ en de ‘NVB Standaard’

DNB heeft als doel organisaties waarop zij toezicht houdt, inzicht te geven in haar beleidspraktijk door bijvoorbeeld een interpretatie van wettelijke vereisten te geven (Q&A), evenals voorbeelden van manieren waarop aan de wettelijke vereisten kan worden voldaan (Good Practices). Het is hierbij belangrijk op te merken dat noch de Q&As, noch de Good Practices van DNB juridisch bindend zijn.

De NVB Standaard beschrijft de technische modeldocumentatie die banken kunnen gebruiken bij de toepassing van advanced analytics, AI en machine learning modellen. Banken bepalen zelf hoe zij de NVB Standaard gebruiken en elke bank bepaalt haar eigen beleid en risk appetite.

Deze NVB Standaard is tot stand gekomen in overleg met De Nederlandsche Bank (DNB), de Autoriteit Financiële Markten (AFM) en de Autoriteit Persoonsgegevens (AP).

Bijlage

1 Model Process Control en Model Governance

Alle onderdelen in deze sectie moeten worden toegelicht vanuit het perspectief van het proces. Alle technische details horen in latere secties thuis.

- 1.1 Model Governance rollen
- 1.2 TMD auteurs
- 1.3 Logboek document geschiedenis
- 1.4 Goedkeuringen en aftekeningen (inclusief gebruik van data)
- 1.5 Ongoing monitoring plan
- 1.6 Model review vereisten

2 Samenvatting

In deze context kan een model verwijzen naar een enkel machine learning of analytics model, een rule-based model of naar een combinatie van machine learning modellen en regels.

2.1 Business problem statement

Beschrijf de samenstelling van het model (machine learning, rule-based of combinatie) en licht het doel van het model toe, d.w.z. op welk business probleem richt het zich, wat wordt het geacht te doen, hoe past het in de bestaande business context en -processen, en wat is de scope voor de klant.

2.2 Beoogd gebruik van het model

Beschrijf in detail de scope van het model en het gebruik ervan in productie. Bijvoorbeeld, hoe vaak wordt output gegenereerd en hoe wordt deze toegepast in processen. Beschrijf alle beoogde toepassingen van het model. Vermeld ook situaties waarin het model niet kan worden gebruikt.

2.3 Data en Modelling

Beschrijf de toepasbaarheid van de gebruikte data en de feitelijke data conditie. Beschrijf het gekozen model, de belangrijkste modelleringsbesluiten en waarom deze keuzes zijn gemaakt.

2.4 Performance

Beschrijf de verwachte impact van het model.

3 Concept modelontwerp

Noem of verwijst naar alle belangrijke beslissingen die samen met business partijen zijn gemaakt en die van invloed zijn geweest op het ontwerp.

3.1 Doelvariabelen definiëren

Beschrijf de doelvariabele (de variabele die we proberen te meten/voorspellen), vermeld op welke business processen deze betrekking heeft en wat het verband is tussen de doelvariabele en de probleemstelling.

3.2 Definitie van een observatie in het model

Beschrijf in business termen een generieke observatie (datapunten / features) die in het model worden gebruikt. Deze beschrijving moet duidelijkheid bieden voor zowel inhoudelijke experts, ontwikkelaars, en validators.

3.3 Scope van het model

Beschrijf de observaties die in scope zijn als het model in productie gebruikt wordt (of gaat worden) en maak daarbij gebruik van relevante business terminologie. Neem in deze beschrijving alle relevante aspecten op, zoals tijd, de definitie van bijv. klant, transactie en/of events die impact kunnen hebben op het model.

4 Technisch modelontwerp en data

4.1 Databronnen

Beschrijf alle gebruikte databronnen, bijv. de database, namen van tabellen/bestanden en externe bronnen (indien relevant). Geef voor elke tabel een lijst van de geëxtraheerde variabelen en hun definities. Beschrijf ook de relatie tussen deze data, het business proces, de probleemstelling en de oplossing. Uit deze lijst moet duidelijk worden hoe het gebruik van data wordt geminimaliseerd, rekening houdend met de GDPR.

4.2 Ophalen van data

Beschrijf vanuit technisch perspectief hoe datastromen tussen IT-systemen lopen en worden geladen om te worden gebruikt voor het doel van het model.

4.3 Beoordeling van datakwaliteit en representativiteit

Beschrijf nauwkeurig de uitgevoerde data kwaliteitschecks en de daaruit voortvloeiende resultaten. Concludeer of deze beoordelingen hebben aangetoond dat de gebruikte data voldoende representatief is voor het gedefinieerde business probleem.

4.4 Dataprocessing

4.4.1 Data pre-processing

Beschrijf de belangrijkste selecties en pre-processing stappen die zijn uitgevoerd, zoals selecties van rijen in tabellen, of anderszins gespecificeerde selecties van ruwe data, omgang met ontbrekende waarden, aggregaties, samenvoegingen, tekst pre-processing, etc. Het ontwerpen en/of definiëren van feature- en/of doelwaarden valt hier niet onder.

4.4.2 Ontwerp feature variabelen

Beschrijf hoe de pre-processed data worden omgezet bij het afleiden van de waardes voor elke feature. Verstrek de input van de business expert die tot deze waardes heeft geleid. De beschrijving van de features moet met beperkte verwijzing naar bestaande code begrijpelijk zijn. Beschrijf ook features die direct uit een database zijn

geselecteerd (bijv. feature store). Lever voor de finale feature set (analytical base table) de voltooidde kwaliteitscontroles, inclusief tests voor bijv. multicollineariteit.

4.4.3 Ontwerp doelvariabele processing

Beschrijf hoe je data hebt verwerkt tot doelwaarden, in overeenstemming met de definitie van de doelvariabele. Je beschrijvingen moeten duidelijk zijn, zodat ze kunnen worden begrepen door niet-technische stakeholders. Zorg dat de juiste documentatie over de exacte keuzes voor de implementatie van doelvariabelen hier of in de base code is opgenomen. Vermeld, indien van toepassing, de timing relatie tussen business termen en doeldefinitie.

5 Modelontwikkeling

5.1 Methodologie

5.1.1 Algemene beschrijving van hier toegepaste machine learning methoden

Beschrijf de verschillende machine learning algoritmes die tijdens het experimenteren zijn gebruikt. Het is belangrijk om een balans te vinden in hoe gedetailleerd dit wordt beschreven versus de details waarnaar verwezen kan worden in bestaande gerenommeerde publicaties.

5.1.2 Train/valideer/test (voor supervised learning modellen)

Beschrijf hoe de trainings/validatie/test sets zijn samengesteld en of er steekproeftechnieken zijn

gebruikt.

5.1.3 Hyperparameter tuning

Beschrijf het hyperparameter tuning proces dat gebruikt is voor het uiteindelijk geselecteerde model.

5.1.4 Domeinexpertise

Vermeld hier, of verwijst naar een ander document waarin alle belangrijke beslispunten staan beschreven die samen met business partijen zijn opgesteld en die impact hebben op het model-ontwerp.

5.2 Validatie van resultaten

5.2.1 Performancemetingen en beperkingen

Geef definities van de performancemetingen die geselecteerd zijn om de geschiktheid van je model te valideren; beschrijf waarom ze relevant waren.

5.2.2 Behaalde resultaten

Verschaf actuele schattingen van de meetwaarden uit de vorige paragraaf voor je uiteindelijke model, evenals een overzicht van de behaalde resultaten voor de modellen die zijn afgewezen. Beschrijf ook de beperkingen van het model die direct voortvloeien uit de meetwaarden.

5.3 Stabiliteitstesten

Stabiliteitstesten worden gebruikt om te bepalen hoe een model in de loop van de tijd verandert.

5.3.1 Beschrijving van toegepaste procedures en statistieken

Beschrijf alle procedures die zijn toegepast om de stabiliteit van je model te testen.

5.3.2 Stabiliteitsschattingen en justification rationale

Geef actuele schattingen van alle overwogen stabiliteitsstatistieken voor je uiteindelijke model. Geef ook een onderbouwing voor het accepteren van de verkregen resultaten. Een belangrijk aspect om te overwegen, is het verkrijgen van bevredigende resultaten die de performance van het model in productie voldoende garanderen.

5.3.3 Gevoeligheidsanalyse

Gevoeligheidsanalyse bepaalt hoe veranderingen in de feature set de doelvariabele kunnen beïnvloeden. Beschrijf alle procedures gebruikt bij het testen van de gevoeligheid van het model.

6 Uitlegbaarheid

Verwijst naar het vermogen om uit te leggen wat er in je model gebeurt van input tot output.

6.1 Beschrijf de vereiste mate van uitlegbaarheid

Beschrijf de toepasselijkheid van uitlegbaarheid met betrekking tot de manier waarop het model wordt gebruikt (bijv. of het model integraal deel uitmaakt van het risicobeheersingsproces en bijgevolg de mate waarin uitlegbaarheid vereist is). Uitlegbaarheid beschrijft de werking en de resultaten van het model voldoende om begrepen

te worden door belanghebbenden. Twee aspecten van uitlegbaarheid zijn transparantie en interpreteerbaarheid. Transparantie verwijst naar het beschrijven van de interne werking van een model, zoals in deze documentatie wordt beschreven. Voor interpreteerbaarheid beschrijft dit hoe een mens de output van het model kan interpreteren.

6.2 Beschrijf hoe interpreteerbaarheid is gewaarborgd (gebaseerd op de vereisten van 6.1)

Beschrijf voor de interpreteerbaarheid (die verwijst naar het begrijpen van de redenering achter de voorspellingen van een model), in het geval modellen zijn ingebed in bestaande processen met interpreteerbaarheidseisen, de lijst van features en hun relatie met het integrity risk framework, indien van toepassing. Beschrijf het belang van de features voor alle steekproeven (globale uitlegbaarheid) en beschrijf de relaties van features met de model output voor specifieke steekproeven (lokale uitlegbaarheid).

7 Risico-overwegingen

Geef een gedetailleerd overzicht van mogelijke risico's en tekortkomingen van het/de toegepaste model/ modellen en technologieën. Onderstaande opsomming is niet bedoeld als uitputtend.

7.1 Regulatory beoordeling

Beschrijf de resultaten van de regulatory self-assessment om aan te tonen dat het model voldoet aan relevante regelgeving en beleid.

7.2 Randvoorwaarden en restricties

Beschrijf de randvoorwaarden en restricties voor het model, bijvoorbeeld 'Data en labelling kwaliteit', 'Beperkingen in scoping' en 'Complexiteit van het model'.

7.3 Aannames

Beschrijf de onderliggende aannames die zijn gemaakt tijdens de ontwikkeling van het model, bijv. 'we veronderstellen dat entiteiten die betrokken zijn bij witwassen zich anders gedragen, of geprofileerd kunnen worden aan de hand van bepaalde kenmerken.'

7.4 Ethische overwegingen

Beschrijf de ethische overwegingen met betrekking tot dit model. Dit kunnen bijv. overwegingen zijn met betrekking tot specifieke ethische onderwerpen (fairness, uitlegbaarheid, gelijkwaardigheid, menselijke beoordeling en bias) op basis van interne en externe regelgeving, beleid en/of richtlijnen. Overwegingen moeten een korte beschrijving van de issues bevatten, zowel data als modellering omvatten, de potentiële impact beoordelen, business overwegingen omvatten en technische keuzes en uitgevoerde analyses documenteren (mogelijk door verwijzingen op te nemen naar specifieke documenten, indien aanwezig).

7.5 Technologisch landschap

Impact van (verwachte) veranderingen in het technologisch landschap die de validiteit van het model in de loop der tijd zouden kunnen beïnvloeden.

7.6 Conclusie

Geef een conclusie over de risico-overwegingen, waarbij de impact op de klanten van financiële instellingen en het doel van het model tegen elkaar worden afgewogen. De toelichting moet de conclusie rechtvaardigen dat eventuele restrisico's tot een aanvaardbaar niveau zijn teruggebracht, in verhouding tot de noodzaak om witwassen adequaat te voorkomen, voor zowel financiële instellingen als de maatschappij.

8 Modelonderhoud in productie

8.1 Monitoring

Describe in detail the plan to monitor the model's performance in production. This should include all relevant model performance metrics as well metrics associated with model and features stability assessment.

Beschrijf in detail het plan om de performance van het model in productie te monitoren. Dit moet alle relevante performance metrics van het model omvatten, evenals meetwaarden die verband houden met de beoordeling van de stabiliteit van het model en de functies.

8.2 Her-training (indien van toepassing)

Beschrijf in detail hoe het model opnieuw wordt getraind als het eenmaal in productie is, inclusief de criteria die aanleiding geven tot het hertrainen van het model, zoals een specifieke tijdlijn of achteruitgang van de performance van het model buiten een bepaalde range.

9 Implementatie van het model

Beschrijf hoe het model in productie zal worden uitgerold. Dit moet alle relevante details bevatten, zoals verwijzing naar het platform, maar ook tests (unit/integratie/functioneel) etc. Beschrijf ook de beperkingen bij het implementeren van het model (vooral in de context van besluitvorming). Besteed, indien relevant, speciale aandacht aan de (geplande) herziening van het model met regelmatige frequentie, d.w.z. wanneer het model (parameters/coëfficiënten) dynamisch wordt geupdate, voorzien van nieuwe data. Als hiervoor een algoritme wordt gebruikt, beschrijf dit algoritme dan duidelijk en de tests die op dit algoritme zijn uitgevoerd.



© Juni 2024

Nederlandse Vereniging van Banken
Gustav Mahlerplein 29-35
1082 MS Amsterdam
www.nvb.nl