

## Belastingdienst/Toeslagen Risicoclassificatie van aanvragen



Utrecht, 28 juni 2013

# Introductie

## ACHTERGROND

- Wens Belastingdienst tot betere en efficiëntere afwikkeling van toeslagen.
- Wens Belastingdienst tot terugdringen van toeslagfraude.
- Wens Belastingdienst tot validatie van een reeds geïmplementeerd model.
- Gezamenlijke workshop 12 april 2013 en follow up 19 april 2013.
- Vanaf april gestart met ontwikkelen en uitvoeren van risicoclassificatiemodel voor huurtoeslag en kinderopvangtoeslag.

## DOEL

- Toelichten en bespreken risicoclassificatie vanaf april.



# Agenda

## Evaluatie Profiling

- Hoe werkt het model (risicoregels versus algoritmen)?
- Wat is tot nu toe bereikt?
- Wat ging goed?
- Wat ging niet goed?
- Lessons learned, hoe verbeteren? (kennis, vaardigheden, hardware, software)



# Evaluatie Profiling

## Doelstelling

- Versneld implementeren en doorontwikkelen van de gehanteerde methodiek van “profiling” van toeslag aanvragers en aanvragen.
- Scoren van aanvragen gedurende een “freeze”.
- Validatie door middel van steekproeven.
- Overdracht kennis aan Belastingdienst.

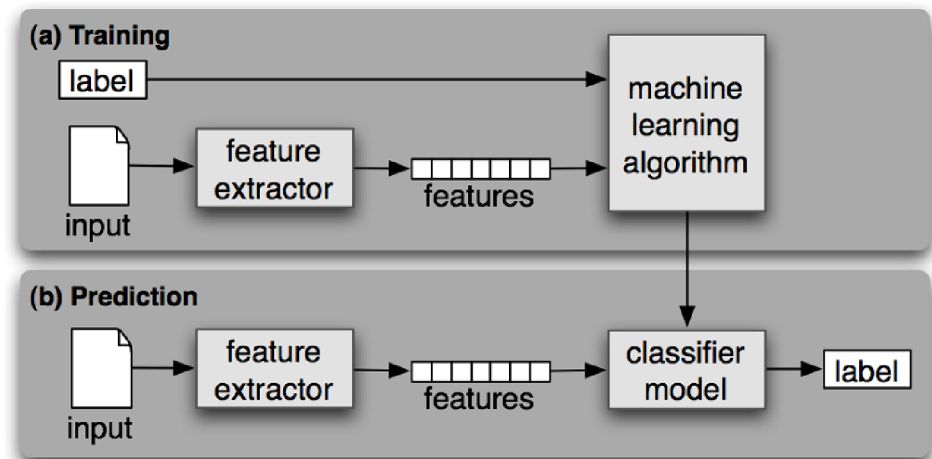


## Lerend model

*Hoe werkt machine learning?*

### Machine learning:

1. Indicatoren definitie
2. Verzamelen contra informatie
3. Verzamelen voorbeeld dossiers
4. Training van het model
5. Scoren van het model



# Machine learning en scoring

## *Aanpak bij Belastingdienst/Toeslagen*

- **Samenstelling van het team:**

Het team verantwoordelijk voor de ontwikkeling en implementatie van het risicoclassificatiemodel bestaat uit medewerkers van de Belastingdienst/Toeslagen ondersteund door specialisten van Deloitte. Voor het ontsluiten en koppelen van databases is tevens ondersteuning verleend door medewerkers van SAS.

- **Aanpak:**

Uit de toeslagaanvragen vloeien gegevens voort (hierna: data). Deze data wordt verder verrijkt en geanalyseerd in een datamodel met referentiebestanden. Dit model vormt de basis van het classificatiemodel.

De referentiebestanden bestaan uit de reeds lopende beschikkingen verkregen van de Belastingdienst Toeslagen en bevatten onder meer de leeftijd van een BSN, inkomensgegevens van de aanvrager, kinderopvang gegevens en adresgegevens van de aanvrager en het kind waarvoor toeslag wordt aangevraagd (bij kinderopvangtoeslag) en huurprijs, rekenhuur en adres (bij huurtoeslag).

- **Gebruikte brongegevens (databases):**

Het model maakt gebruik van variabelen afkomstig uit een aantal databases:

- BVR- database;
- PIT- database;
- Datamart (conceptbeschikkingen en lopende toeslagen);
- FRS- database;
- COA- database.
- TSL- database.



## De aanpak in de praktijk

### *Model werkzaamheden Belastingdienst/Toeslagen*

| Activiteit                    |                                                                                                                                                                       | Frequentie                                                                |
|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------|
| Indicatoren definitie         | Met experts van Toeslagen zitten om te bepalen waar een mogelijke indicator van fraude zou kunnen zitten.                                                             | 1 maal per kwartaal.                                                      |
| Verzamelen contra informatie  | Opvragen van informatie die nodig is om een indicator te formuleren.                                                                                                  | Afhankelijk van soort informatie variërend van maandelijks tot jaarlijks. |
| Verzamelen voorbeeld dossiers | Verzamelen van 20% foute en 80% goede beschikkingen. Minimaal 1500 tot 2000 dossiers nodig.                                                                           | Elke 1 of 2 maanden.                                                      |
| Training van het model        | Aan de hand van de voorbeeld dossiers het model een profiel van goede en foute beschikkingen leren.                                                                   | Elke 1 of 2 maanden.                                                      |
| Scoren van het model          | Aan de hand van het model de nieuwe beschikkingen scoren. De score is de gelijkenis met een foute beschikking (score 1,00) of met een goede beschikking (score 0,00). | Elke maand tijdens de freeze.                                             |
| Afhandeling                   | Aan de hand van de scores (risico – niet rechtmatigheid) bepalen hoe de dossiers behandeld moeten worden.                                                             | Elke maand tijdens de freeze.                                             |

## Resultaten scoring april, mei, juni 2013

| Maand | Wat is er gescoord?                                                              | Resultaat scoring                                                                                                                     | Gevolgprocedure                                                                                                                                     |
|-------|----------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| April | Huurtoeslag, nieuwe conceptbeschikkingen.                                        | 80.712 BSN nummers zijn gescoord met zowel het lerend model als met separate indicatoren. Top 5000 met hoogst risico opgeleverd.      | 4.106 aanvragen over de maand april 2013 worden opgeschort.<br>De beschikking en betaling van 894 aanvragen wordt per direct stopgezet.             |
| Mei   | Kinderopvangtoeslag, lopende beschikkingen en nieuwe conceptbeschikkingen.       | 360.144 lopende beschikkingen en 6.446 nieuwe aanvragen gescoord. 8 "buckets" met risico score opgeleverd en een top 1000 opgeleverd. | Buckets blijken niet goed te werken, daarom handmatig een klein aantal aanvragen uitgesloten van beschikken.                                        |
| Juni  | Kinderopvangtoeslag, nieuwe beschikkingen en mutaties met een toeslag verhoging. | Totaal 428.000 beschikkingen gescoord. Waarvan 14.713 nieuwe concept beschikkingen en mutaties met een hoger toeslag bedrag gescoord. | Top 200 helemaal uitsluiten. Top 200 tot 600 uitsluiten voor zover jaarbedrag boven 15.000. Daarnaast 91 gestopt op grond van "knock out criteria". |
| Juni  | Huurtoeslag, nieuwe beschikkingen en mutaties met een toeslag verhoging.         | Totaal 1.25 miljoen beschikkingen gescoord. Waarvan 21.034 concept beschikkingen met mutaties met een hoger toeslag bedrag.           | Top 200 helemaal uitsluiten 79 uitsluiten op grond van "knock out criteria".                                                                        |

# Ervaringen met scoring en model

## *Overzicht van de evolutie van het model tot juni 2013*

- Hybride aanpak (scoren zowel met indicatoren als met het lerend model) gehanteerd.
- Ontwikkeling en koppeling van halfproducten gebaseerd op de databases.
- Training van het model voor huurtoeslag en kinderopvangtoeslag.
- Aandacht voor controle en datakwaliteit.
- Hardware en software omgeving.
- Aandacht voor gevolgen en implementatie in het werkproces.



# Ervaringen met scoring en model

## Overzicht van de evolutie van het model tot juni 2013

### April – Scoring huurtoeslag

- Hybride aanpak (scoren zowel met indicatoren als met het lerend model) gehanteerd. Dit omdat het lastig bleek voldoende gewenste “foute dossiers” op korte termijn te verkrijgen om het model te trainen. Ook “gedeeltelijk fout” dossiers hadden een negatief effect op de werking van het model.
- Met de indicatoren kon worden gefilterd op onrechtmatigheden in aanvragen, zodoende kon een hoog risico aanvraag tevens los van het lerend model worden geïdentificeerd.
- Datakwaliteit bleek een aandachtspunt. Zo bleek bijvoorbeeld naderhand een indicator “BSN-leeftijd” niet correct te zijn samengesteld in de data, hierdoor werd de indicator “BSN-leeftijd” minder betrouwbaar.
- Het model is gebouwd in een SAS omgeving, hiertoe is tevens ondersteuning geboden door SAS consultants bij het verwerken van de scoring van het model.
- Hardware, omvang van de databases en doorlooptijd bleken vertragende factoren. Uiteindelijk is besloten een van de grote databases (BAG-lijst) niet op te nemen in het model omdat het bestand te groot was om op de beschikbare hardware te kunnen bewerken binnen de beschikbare tijd.
- Na evaluatie bleken sommige als indicator gebruikte databronnen minder betrouwbaar (bijvoorbeeld PIT-database en de “blacklist”) waardoor “false positives” ontstonden.
- De aansluiting tussen de scoring van het model en de workflow van de Belastingdienst kon in een volgende fase verbeterd worden.



# Ervaringen met scoring en model

## *Overzicht van de evolutie van het model tot juni 2013*

### **Mei – Scoring kinderopvangtoeslag**

- In mei voortbordurend op de hybride aanpak van april is een lijst met indicatoren samen met de experts van de Belastingdienst (deels) herzien en zijn nieuwe indicatoren geformuleerd. Dit resulteerde in +/- 40 indicatoren voor kinderopvangtoeslag.
- De gebruikte databronnen van april zijn geëvalueerd en waar nodig herzien. Zo is PIT niet meer gebruikt als indicator en is de BSN-leeftijd opnieuw opgesteld aan de hand van nieuwe data
- Gekozen is om na scoring van alle aanvragen, deze in te delen in “buckets” gebaseerd op een indicator (bijvoorbeeld “ouder-kind relatie”).
- Het verkrijgen van voldoende trainingsdossiers bleek wederom een uitdaging. Er is ter training gebruik gemaakt van voorhanden zijnde bekende dossiers, gescoord door experts van de Belastingdienst. Hierdoor bleek het lerend model eenzijdig getraind te worden, dit had “false positives” tot gevolg.
- Omdat hardware en software (SAS omgeving) een vertragende werking hadden in april is er in mei voor gekozen om een SAS consultant (en SAS software + bijbehorende hardware) in te zetten op het project.
- Na scoring bleek dat opdeling in “buckets” niet het gewenste resultaat had. Daarom is er bij besluitvorming omtrent de afhandeling van de aanvragen gebruik gemaakt van een scoringlijst op totaal score niveau (van hoog naar laag).

# Ervaringen met scoring en model

## *Overzicht van de evolutie van het model tot juni 2013*

### **Juni – Scoring huurtoeslag en kinderopvangtoeslag**

- Het kinderopvangtoeslag model is niet opnieuw getraind. Ook de indicatoren zijn in stand gehouden. Voor huurtoeslag zijn extra indicatoren opgesteld, hiermee komt het totaal op +/- 60 indicatoren voor kinderopvangtoeslag en huurtoeslag.
- De volledige data gebruikt voor zowel kinderopvangtoeslag als huurtoeslag is geüpdatet om recente informatie te gebruiken.
- Een aantal van de vastgestelde indicatoren voor het huurtoeslag model zijn komen te vervallen.
- Ook in juni is een hybride aanpak gehanteerd. Met de indicatoren kon worden gefilterd op onrechtmatigheden. Op deze manier is in juni tevens gefilterd op onrechtmatigheden (bijvoorbeeld aanvragers buiten Nederland) van aanvragen binnen de gescoorde populatie.
- Het verkrijgen van voldoende trainingsdossiers voor de huurtoeslag bleek wederom lastig. Er was nog geen terug voeding van de aanvragen gescoord in april om het model mee te trainen.
- Er is ter training gebruik gemaakt van voorhanden zijnde bekende dossiers, gescoord door experts van de Belastingdienst. Het betrof dossiers met voornamelijk vastgestelde fraudegevallen uit voorgaande acties. Omdat deze een relatief identiek fraude patroon hebben is dit van invloed geweest op de training van het model.

## Lessons learned

- Hybride aanpak (scoren zowel met indicatoren als met het lerend model) heeft de voorkeur.
- Risico scoring is afhankelijk van de voorbeeld dossiers en training. Terugvoeding van resultaten van gescoorde aanvragen aan het model draagt bij aan training.
- Een niet representatieve training set – is van invloed op het model.
- Gedeeltelijk compliante dossiers aanbieden (“een beetje fout”) maakt het model slechter, niet beter.
- Het model geeft alleen een risico score af; geen oordeel over compliantie.
- De resultaten van de scoring moeten goed aansluiten op de workflow.
- Hardware en software is van invloed op de doorlooptijd van de scoring.
- Inhoudelijke informatie over het model en de werking wordt idealiter verstrekt op een “need to know” basis.
- Datakwaliteit en datacontrole is een punt van aandacht.



## Vervolg stappen

*Planning tot en met oktober 2013*



- Evaluatie resultaten
- Bepalen update periodes
- Training
- Toevoegen
- Documentatie

Toevoegen  
Training  
Versie beheer  
Documentatie  
Scoring  
Updates HPs (high level)

- Toevoegen
- Training
- Versie beheer
- Documentatie
- Scoring
- Updates HPs (high level)

Zelfstandig runnen en beheersen

- Zwaartepunt op training medewerkers;
- Versterkte focus op documenteren van processtappen;
- Aandacht voor terug voeden van resultaten aan het model;
- Aandacht voor kwaliteit, controle en verantwoording.



# Steekproeven

## *Controle, beheersing en groei*

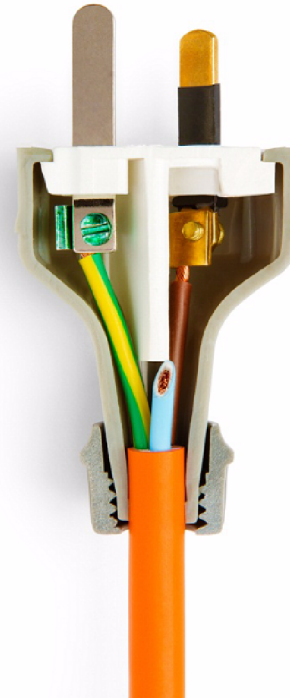
Waarom zijn steekproeven nodig?

**1) Het lerende model moet gevoed worden met nieuwe gevallen.**

Willekeurige selectie van onbekende dossiers. Dit voedt het model met onbekende gevallen met mogelijk nieuwe onrechtmatigheid-patronen

**2) Validatie van resultaten.**

Beheerste en beperkte controle op de resultaten van de scoring door het model. Zorgt voor gerechtvaardigd vertrouwen in de afhandeling van aanvragen.



# Vragen?



# Contactgegevens

|                                          |             |
|------------------------------------------|-------------|
| <div></div>                              |             |
| <div></div>                              | <div></div> |
| Deloitte Risk Services                   |             |
| Tel: +31 <div></div>                     |             |
| <div></div> <a href="#">@deloitte.nl</a> |             |

|                                          |             |
|------------------------------------------|-------------|
| <div></div>                              |             |
| <div></div>                              | <div></div> |
| Deloitte Risk Services                   |             |
| Tel: +31 <div></div>                     |             |
| <div></div> <a href="#">@deloitte.nl</a> |             |

|                                          |             |
|------------------------------------------|-------------|
| <div></div>                              |             |
| <div></div>                              | <div></div> |
| Deloitte Risk Services                   |             |
| Tel: +31 <div></div>                     |             |
| <div></div> <a href="#">@deloitte.nl</a> |             |

|                                          |             |
|------------------------------------------|-------------|
| <div></div>                              |             |
| <div></div>                              | <div></div> |
| Deloitte Risk Services                   |             |
| Tel: +31 <div></div>                     |             |
| <div></div> <a href="#">@deloitte.nl</a> |             |



Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited, a UK private company limited by guarantee, and its network of member firms, each of which is a legally separate and independent entity. Please see [www.deloitte.nl/about](http://www.deloitte.nl/about) for a detailed description of the legal structure of Deloitte Touche Tohmatsu Limited and its member firms.

Deloitte provides audit, tax, consulting, and financial advisory services to public and private clients spanning multiple industries. With a globally connected network of member firms in more than 150 countries, Deloitte brings world-class capabilities and high-quality service to clients, delivering the insights they need to address their most complex business challenges. Deloitte has in the region of 200,000 professionals, all committed to becoming the standard of excellence.

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms, or their related entities (collectively, the "Deloitte Network") is, by means of this publication, rendering professional advice or services. Before making any decision or taking any action that may affect your finances or your business, you should consult a qualified professional adviser. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person who relies on this communication.