

# PINNs: de gouden combinatie van natuurkundig en datagedreven modelleren



Foto: Rauf Karimov

De relatief nieuwe *Physics-Informed Neural Network*-modellen bundelen de krachten van traditionele ‘verkeersstroomtheorie’-modellen en van datagedreven machine-learningmodellen. Hoe werken ze? Welke toepassingen zijn denkbaar? Waar wordt nog aan gewerkt? In deze nieuwe AI-tutorial bespreken Ying-Chuan Ni en Theivaprakasham Hari wat het werkveld heeft aan deze zogenaamde PINNs.

## Het waarom en hoe van PINNs

Wie ooit een opleiding in de verkeerskunde heeft gevolgd, zal er-mee bekend zijn: het *Lighthill-Whitham-Richards-model*, kortweg LWR-model. Dit model benadert voertuigen op een weg als een stromende vloeistof. Met één enkele natuurkundige vergelijking<sup>1</sup> verklaart het hoe verkeer zich gedraagt – waarom bijvoorbeeld een kleine vertraging aan de kop van een file zich als een golf stroomopwaarts kan voortplanten en bestuurders kan bereiken die zich kilometers van de oorspronkelijke verstoring bevinden.

Deze *natuurkundige insteek* heeft ons decennialang geholpen om verschijnselen als congestie te bestuderen en doorgronden. Toch was

ook al snel duidelijk dat in het verkeer de natuurkundige vlieger nooit helemaal opgaat. Verkeer is namelijk rommeliger dan een vloeistof. Zaken als verschillen in rijstijl, afleiding, weersomstandigheden enzovoort zorgen ervoor dat het verkeer geregeld anders stroomt dan je op basis van de natuurkundige vergelijking zou verwachten.

Verkeerskundigen waren dan ook wat blij toen het *datagedreven machine learning* z'n intrede deed. Met deze aanpak voeden we computers met data van sensoren, verkeerscamera's en mobiele telefoondata en laten we de computer de onderliggende patronen op een netwerk identificeren. Die patronen zijn weer de basis voor voorspellingen over dat netwerk.

En het moet gezegd dat die datagedreven modellen zeer nauwkeurig voorspellen – veel nauwkeuriger dan met ‘natuurkundige modellen’ mogelijk is. Toch is ook hier een maar: de omstandigheden waarvoor je het model laat voorspellen, moeten wel vergelijkbaar zijn met de historische data waarop ze zijn getraind. Betreffen die data bijvoor-

<sup>1</sup> Het betreft de vergelijking:  $\partial k / \partial t + \partial q / \partial x = 0$ , waarbij  $k$  staat voor verkeersdichtheid en  $q$  voor verkeersstroom. De vergelijking koppelt veranderingen in verkeersdichtheid (in de tijd) aan veranderingen in verkeersstroom (langs de weg) en drukt daarmee de *behoudswet* van verkeersstromen uit.

beeld reguliere spitsen en laat je het model vervolgens rekenen aan een ongeval, dan kan het rare fouten maken. In het model verschijnen dan ineens voertuigen uit het niets, ze verdwijnen juist van het ene op het andere moment, of er rijden ergens twee keer meer auto's dan de maximale wegcapaciteit ter plaatse. Het probleem is dat de datagedreven machine-learningmodellen wel patronen herkennen in de aangeboden data, maar nul kennis hebben over de natuurkundige wetten en regels van het verkeer. Buiten hun vertrouwde (getrainde) omgeving gokken ze dus maar wat.

### Krachten bundelen

Nu zou je de training van deze datagedreven modellen enorm kunnen verbreden, door ze ook te voeden met minder reguliere situaties, zoals ongevallen, pechgevallen, wegwerkzaamheden enzovoort. Dat gebeurt ook wel, maar die route is behalve kostbaar en tijdrovend – waar haal je al die data vandaan? – ook niet sluitend. Er kunnen zich tenslotte altijd nieuwe omstandigheden voordoen.

In 2017 kwamen Raissi, Perdikaris en Karniadakis daarom met een veel elegantere oplossing: zij combineerden de natuurkundige en de datagedreven benadering. Vandaar ook de naam van hun model: het **Physics-Informed Neural Network** of PINN.

Het *Neural Network* slaat op een specifiek type machine learning, geïnspireerd op de manier waarop neuronen in het menselijk brein werken. Zo'n neuraal netwerk is al geschikter om buiten de trainingsdata om te redeneren. Stel je zo'n netwerk bloot aan een groot aantal voorbeelden, dan past het geleidelijk zijn interne parameters aan tot dat het nauwkeurige voorspellingen kan doen voor óók (aanpalende) situaties die het nog niet kent.

Maar de echte vernieuwing van PINN zit in het *Physics-Informed* in de naam. In het trainingsproces van het netwerk worden natuurkundige wetten namelijk expliciet meegenomen. Bij elke voorspelling die het model daarna doet, wordt ook altijd op twee punten gecontroleerd: blijft de voorspelling in lijn met de werkelijke metingen én blijft ze in overeenstemming met de natuurkundige wetten waarop het model is gebaseerd? Wanneer de voorspelling in strijd is met een van die vergelijkingen, worden de modelparameters aangepast.

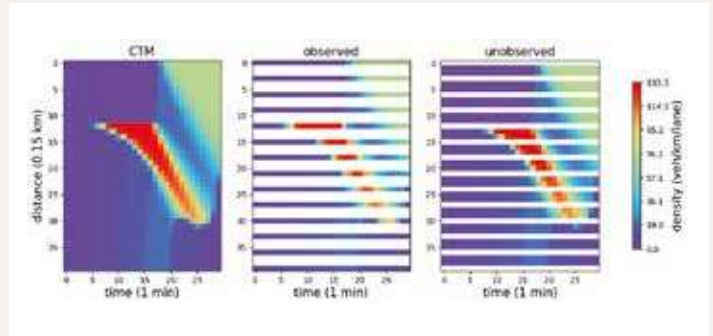
Met die aanpak worden de rare fouten van de gewone datagedreven aanpak voorkomen. Wanneer een PINN nog nooit een specifieke verkeersafwijking heeft waargenomen, 'weet' het in ieder geval dat auto's niet uit het niets kunnen verschijnen of verdwijnen en dat verkeersschokgolven zich moeten voortplanten met snelheden die voldoen aan natuurkundige wetten.

Dit natuurkundige bewustzijn geeft PINNs een groot voordeel op het gebied van generaliseerbaarheid en betrouwbaarheid ten opzichte van gewone machine-learningssystemen en zelfs van de specifiekere neurale netwerken.

### Toepassingen van PINNs

Voor verkeerskundigen ligt de praktische aantrekkingskracht van PINNs vooral in de sterk verminderde afhankelijkheid van enorme datasets. De ingebouwde natuurkundige kennis vult immers op natuurlijke wijze data-lacunes op.

Het is daarom niet verrassend dat PINNs sinds hun introductie zijn toegepast op een breed scala aan vraagstukken binnen verkeersmodelleren en verkeersregeling. Laten we nu bekijken hoe deze techniek ons kan helpen op twee verschillende aggregatieniveaus.



**Figuur 1:**

Contourdiagram van de verkeersdichtheid op een autosnelwegvak van 6 km gedurende 30 minuten.

### PINNs voor verkeersstromen

Door technologische beperkingen en financiële overwegingen zijn verkeersgegevens die door moderne sensoren worden verzameld vaak schaars en ruisgevoelig. *Traffic State Estimation*, TSE, is een proces waarbij de verkeerssituatie in ruimte en tijd op wegvakken wordt afgeleid uit die beperkte (gedeeltelijke) gegevens.

Traditionele TSE-benaderingen zijn ofwel 'natuurkundig' (zoals Kalman-filtering en particle filtering) ofwel datagedreven (lineaire regressie, Bayesiaanse netwerken enzovoort). We hadden al besproken dat die aanpakken hun beperkingen hebben. Hoe pakken PINNs een TSE aan?

Stel dat op een autosnelweg slechts om de 450 meter lusedetectoren aanwezig zijn. Halverwege het traject vindt een ongeval plaats, waardoor een file ontstaat die zich stroomopwaarts voortplant – zie figuur 1. De detectoren registreren alleen wat er op hun eigen locatie gebeurt (middelste diagram).

Een PINN gebruikt de beschikbare meetgegevens én een verkeersstroommodel om de ontbrekende verkeersdichtheden tussen de detectoren te reconstrueren (rechterdiagram). Daardoor ontstaat een veel completer beeld van de verkeerssituatie dan op basis van de metingen alleen.

Bij een traditioneel neuraal netwerk wordt uitsluitend gekeken hoe goed de voorspellingen overeenkomen met de beschikbare data. Bij een PINN wordt daarnaast voortdurend gecontroleerd of de voorspellingen ook voldoen aan de onderliggende verkeerskundige theorie. De verliesfunctie (de 'voorspelfout' die correctie behoeft) bestaat daarom uit een dataterm  $L_{data}$  én een natuurkundige term  $L_{physics}$ :

$$L_{total} = \omega_d L_{data} + \omega_p L_{physics}$$

Aan beide termen wordt een wegingsfactor  $\omega$  toegekend.<sup>2</sup>

Voor de natuurkundige term kan een vergelijking als de behoudswet van het LWR-model worden gebruikt, maar ook een geavanceerder tweede-orde verkeersstroommodel, zoals het Aw-Rascle-Zhang-model, dat *stop-and-go*-golven beter beschrijft.

Door de historische data zo te toetsen aan verkeerskundige wetten heeft een PINN minder trainingsdata nodig, wordt voorkomen dat het

<sup>2</sup> Daarmee geeft de modelontwikkelaar aan hoe zwaar de meetgegevens dan wel de natuurkundige vergelijking meewegen bij de foutbepaling. Die verhouding kan bijvoorbeeld 0,7 voor data ( $d$ ) zijn, tegenover 0,3 voor de natuurkundige term ( $p$ ), of juist 0,1 tegenover 0,9.

model zich té sterk aanpast aan trainingsdata ('overfitting') en kan het ook beter omgaan met ruis in meetgegevens.

Dezelfde aanpak kan ook worden gebruikt voor verkeersvoorspellingen. Juist wanneer gegevens onvolledig zijn of sensoren tijdelijk uitvallen, kan de toevoeging van natuurkundige kennis de voorspellende prestaties aanzienlijk verbeteren.

### PINNs voor voertuigbewegingen

PINNs kunnen ook worden toegepast op het gedrag van individuele voertuigen.

Neem volggedrag. Onderzoekers combineren hiervoor een deep-learningmodel met een zogenoemd *car-following model*, zoals het Intelligent Driver Model (IDM), het Optimal Velocity Model of het Wiedemann-model. Deze modellen beschrijven hoe bestuurders hun snelheid aanpassen aan die van hun voorligger. Door op deze wijze verkeerskundige kennis te combineren met machine learning kunnen voertuigbewegingen worden gereconstrueerd of zelfs voorspeld.

Figuur 2 toont een voorbeeld van een PINN-architectuur voor de voorspelling van longitudinale voertuigtrajecten. Het model gebruikt onder meer de huidige snelheid, de volgafstand en het snelheidsverschil met de voorligger om de volgende versnelling te berekenen. Tijdens het trainen wordt de voorspelde voertuigpositie niet alleen vergeleken met historische gegevens, maar ook getoetst aan de natuurkundige beperkingen die in het volgmodel zijn vastgelegd. De totale verliesfunctie bestaat daarom opnieuw uit een dataterm ( $L_{data}$ ) en een natuurkundige term ( $L_{physics}$ ).

Naast volggedrag is er ook nog de uitdaging van de rijstrookwisselingen. Dat is een lastige, want zodra voertuigen niet alleen in de lengterichting maar ook zijdelings bewegen, neemt de complexiteit sterk toe. Maar ook hier – juist hier – zouden PINNs kunnen helpen. De verwachting is dat de wisselingen beter voorspeld kunnen worden als we empirisch waargenomen rijgedrag combineren met natuurkundige beperkingen van voertuigdynamica.

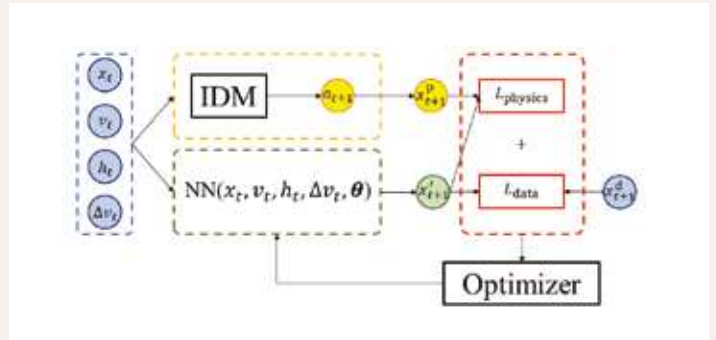
Dat is niet alleen interessant voor microscopische verkeerssimulaties. Ook voor voertuigen met geavanceerde rijhulpsystemen of autonome rijfuncties wordt het steeds belangrijker om het gedrag van andere weggebruikers correct in te schatten. PINNs kunnen daarom bijdragen aan een veiligere, efficiëntere en betrouwbaardere verkeersomgeving waarin menselijke bestuurders en geautomatiseerde voertuigen samen deelnemen aan het verkeer.

### De beperkingen van PINNs

Genoeg nieuwe kansen dus met PINNs. Toch willen we de verwachtingen een beetje temperen: ook de 'gouden combinatie' van natuurkundig en datagedreven heeft namelijk z'n beperkingen.

Om te beginnen is de natuurkundige correctie zo goed als de onderliggende vergelijkingen. Wanneer ontwikkelaars het algoritme voorzien van bijvoorbeeld te eenvoudige vergelijkingen die geen rekening houden met wegdekcondities, reactietijden van bestuurders of de effecten van extreme weersomstandigheden, zal het neurale netwerk juist die onjuiste voorstelling van de werkelijkheid afdwingen. De intelligentie van een PINN wordt dus begrensd door de kwaliteit van de natuurkundige (wiskundige) waarheid die wij eraan meegeven.

Daarnaast vormt de optimalisatie van deze architecturen een uitdaging. Zoals de eerder genoemde verliesfunctie laat zien, probeert het



**Figuur 2:**  
Opzet van een PINN voor de voorspelling van voertuigtrajecten.

systeem voortdurend twee verschillende doelen tegelijk te dienen. Enerzijds moet het overeenkomen met de daadwerkelijke verkeerswaarnemingen, anderzijds moet het voldoen aan de opgelegde natuurkundige randvoorwaarden. Het vinden van de juiste balans tussen het gewicht van de dataterm en het gewicht van de natuurkundige term vereist vaak veel experimenteren.

Een ander punt is dat het berekenen van de natuurkundige verliesfunctie complexe gradiëntberekeningen vereist. Daardoor nemen de benodigde rekenkracht en trainingstijd aanzienlijk toe ten opzichte van conventionele deep-learningmethoden.

Een laatste kanttekening: hoewel PINNs juist bedoeld zijn om met beperkte hoeveelheden sensorgegevens te werken, zijn ze nog steeds afhankelijk van hoogwaardige data voor een grondige validatie. Voor onderzoek naar microscopische voertuigtrajecten kunnen we gelukkig putten uit omvangrijke open-source datasets. Maar voor macroscopische toepassingen ligt dat anders. Als referentiegegevens zijn gedetailleerde, foutloze dichtheidskaarten met een hoge ruimtelijke resolutie voor grote geografische gebieden nodig – en die zijn voor vooral stedelijke verkeersnetwerken nog onvoldoende voorhanden.

### Conclusie

Zoals we hebben gezien zijn *Physics-Informed Neural Networks* geen vervanging van bestaande verkeersmodellen of van datagedreven machine learning. Het is juist een *combinatie* van beide: door natuurkundige kennis direct in het leerproces op te nemen, kunnen PINNs ook met onvolledige of ruisgevoelige data betrouwbare voorspellingen doen en beter generaliseren naar situaties die niet eerder in de trainingsdata voorkwamen.

Natuurlijk zijn er ook de genoemde uitdagingen. Maar het is het waard die uitdagingen aan te gaan en stappen te maken met PINNs. In ons vakgebied beschrijven we al tientallen jaren verkeersstromen met behulp van natuurkundige modellen op zowel macro- als microniveau. Met de datagedreven aanpak waren we daar weer even los van – maar met de hybride aanpak van PINN kunnen we datagedreven werken én dankbaar gebruikmaken van die decennia aan natuurkundige kennis en modellen. Of het nu gaat om verkeersschattingen, verkeersvoorspellingen, microscopische simulaties of ondersteuning van autonome voertuigen: we kunnen weer wat stappen vooruit zetten. ●

### De auteur

Ying-Chuan Ni is onderzoeker aan de Traffic Engineering Group van ETH Zurich.

Theivaprakasham Hari is onderzoeker van TU Delft's DAIMOND Lab.