

Go to www.menti.com and use the code 9089 4492



**Think of a moment when you had to
communicate data science/ML work to
someone not in your role**

Go to www.menti.com and use the code 9089 4492



What was the role of the person you had to communicate with?

Go to www.menti.com and use the code 9089 4492

What did you have to communicate?



Press ENTER to pause scroll

An introduction to communicating about your models

Dr Cynthia C. S. Liem MMus

Associate Professor, Multimedia Computing Group, Delft University of Technology

c.c.s.liem@tudelft.nl |  @informusiccs

Agenda

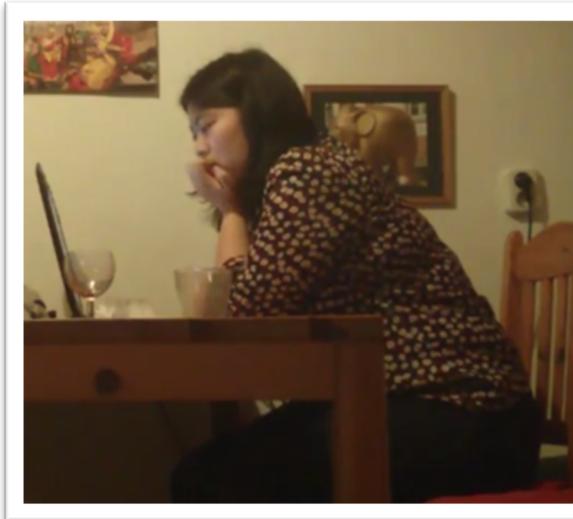
- How I became a communicator
- Communicating interdisciplinarily
- Communicating for accountability
- Communicating to laypeople
- Conclusion

How I became a communicator

My background



© Marco Borggreve



Today



<https://www.youtube.com/watch?v=tnBcVwcoMYY>

<https://www.nrc.nl/nieuws/2021/03/28/als-iets-onvindbaar-is-bestaat-het-niet-a4037549>

INTERVIEW CYNTHIA LIEM INFORMATICUS

'Als iets onvindbaar is, bestaat het niet'

Cynthia Liem is docent informatica in Delft én pianist. 'Dat ik twee expertises heb, heeft me aan beide kanten geholpen.'

Door onze redacteur

Laura Wismans

Onderwijs gaat voor, daar moet je bij aanwezig zijn. Daarom heeft Cynthia Liem haar agenda. Cynthia Liem (34) is universitair docent informatica aan de Technische Universiteit Delft en klassiek concertpianist. 'De hoofdfocus van mijn onderwijs is dat studenten op een dag. Dan moet je wel zorgen dat je de energie hebt om die dag door te gaan. Daarom weet ik waar ze geïntaleert als lid van De Jonge Academie, waar ze zich hard wil maken voor digitale geletterdheid.' Dat ik twee expertises heb heeft me aan beide kanten geholpen', zegt Liem. 'Als pianist wilde ik graag online muzieklessen gaan ontwikkelen, maar dat liep niet goed af. We weten dat mensen niet iedereen snel goed werken met dingen die leren al kent. Ze zijn georganiseerd op artiesten, maar niet op de techniek. In de studie, ik speel werken van componisten die meerdele delen hebben. Als er een probleem is, dan is dat altijd verschillende alegro's. Dat speelt ook bij jazz en wereldmuziek. De stamlijnen en de dansen, alles heel heterotopisch.'

Veel mensen zijn „bij genoeg“ met wat ze vinden. Maar het blijft niet wachten tot dat je dat ook gemakkelijk te vinden is, zag Liem. 'Wat niet te vinden is, bestaat niet. Als promovendus ben ik daarvan overtuigd. Ik veel onderzoek naar algorithmen in de muziek. „Ik heb voor de muziek bestudeerd hoe dat kan.“ De muziek die ik niet heeft me ook bij een breder punt in de kunstmatige intelligentie en zoektoepassingen kunnen de toekomstbaarheid van de systemen. Doen ze wat we denken dat ze doen? Sommige hokjes hebben een effectief acherhaal. Dat was de conclusie die ik in die indelen, heel veel gevonden.'

Ziet een musicus dat eerder dan een informaticus?

„Als musicus ben ik daarom ik wel gevoelens en intuïtie. Informaticus juist exact opgeleid, de belangrijkste fundamenteel zijn wiskunde en logica. In de studie is dat een voorloper. Daar wordt een vaag, menselijk probleem in data geocdeer en dan ga je ermee verder."

Zodra het in cijfers gestopt is, ziet het er objectief uit. De zogenoemde rationaliteit. Informaticus kan dus discriminerende systemen die je nu weet hoeft, hebben daarmee te maken. Uit de data komen de regels. En daarvoor ga je een voorloper. Echter, dat is niet exact opgeleid, misschien zijn bepaalde soorten informatie niet goed codereerbaar. Ik ben niet van mening dat musici sneller kritisch over zijn werk zijn. „Een voorinformaticus komt daardoor niet direct in contact met de wereld. Het was het vooral de roep van dat domine om gegevens te maken. Daarom ben ik niet zo enthousiast over dat ik dat ga in een ander domein kunnen wel data genereren, maar ja dat doen dat niet moet. Daarom ben ik niet zo enthousiast over dat ik een vertaalbaar gelijk samenvoering. Die holistische kijk op informatica vind ik heel interessant."

FOTO: ROBERT DE GROOT

Cynthia Liem: „Ik vind verbinding maken belangrijker dan dat ik een erkend specialist op één vlak word.“

Hoe kan je vrouw in dit vakgebied helpen? *„Of dat moet mijn voorname blik?“*

„Ik hoor vaaker vrouwen in dit vakgebied te spreken. Of dat moet zijn of dat moet mijn voorname blik?“ Dat is niet alleen vrouwen die zo denken, ik denk wel dat het voor vrouwen misschien nog iets belangrijker is om te kunnen bewijzen dat het mij ook dat veel vrouwen in het veld verhinderen proberen te maken dat ze niet deel kunnen nemen. Daarom valt ze wel tussen wal en schip, want ze affilieren zich dus veel minder met de muziek dan de mannen. Ik werk als universitair docent informatica aan de TU Delft en ben lid van De Jonge Academie. Ik speel voor de Magna Duo. Op 23 maart is ze geïnstalleerd als lid van de Jonge Academie.

ding maken belangrijker dan dat ik een erkend specialist op één vlak word.“

Heeft deze aanpak de tijdsgeest mee? *„Weke zijn niet meer met toy problemen bezig.“*

„Beter. We zijn niet meer met toy problemen bezig. Eerder gingen data sets en algoritmen over de voorhoede van lengtes van bloemblaadjes, nu gaat het over wie wie selecteren of wat wij prijzen. Dat is een belangrijke ontwikkeling. Overhoop hangt over het effect van algoritmes op de samenleving. Daarover moet worden gesproken. Daarover moeten mensen moeite doen om recepties te accepteren. Daar zit steeds meer beweging in.“

Hoe ziet dat eruit?

„Ik werkte een tijdje achteraan een algemeen voorlichtingsprogramma van de universiteit. We weten dat een mens subjectief en in mogelijk voorvergaan in de taak, het idee is dat als dat niet meer het geval is, dan zijn algoritmen zijn expliciete, uitgeschreven en systematische proceduren. Daarom is het belangrijk dat recepties dat ook goed kunnen begrijpen. Dan kan jou het recept van mijn favoriete appetitaat geven. Als ik je vertel dat ik een recept heb, dan ga je precies die taart. Maar is het recept daarmee objectief? Vind jij die taart dan ook lekker? Dat is een voorzichtige en exact herhaalbaar, maar of het de beste procedure is, is de vraag.“

Mensen kunnen bij recepten heel goed tussen de regels lezen, je doet er misschien niet veel van maar misschien wel. Dat is de crème van recepten. Beide regels bepaalt welke regels hij aanhoudt en welche niet, en een verantwoording alle regels. Je kunt dat niet goed goede redenen hebben om een regel niet te volgen, of om juist wel preteert dat dat een goede reden is. Bij vraagtekens als sollicitaties belangrijk, maar dat is wat we neigen te vergaten: we zijn algoritmes denken."

Wat is dan de verhouding tussen mens en algoritme?

„In een wereld met algoritmen gaat het er niet meer om hoe we efficiënt kunnen handelen en niet meer na hoeven te denken. Een algoritme moet mensen helpen bij het bewoogpatroon dat ze willen.“ De mens moet zich voorbereiden met algoritmes bemoeien, bij het maken van keuzes. „Als dat niet gebeurt, dan gebeurt dat dat versus-denken“ waar de mens niet meer in betrokken is. Ze weken naast elkaar uit."

Hoe kan je dat verbeteren?

„In de psychologie wordt veel gekeken naar validiteit, maar de vraag of verwachte inschattingen dezelfde uitkomst leveren. In de informatica kan meegenomen worden of verwachte systemen dezelfde uitkomst leveren. Ik test dat door de verschillende data op basis van een audiobestand een generatieve classificatie aan de audio-applicatie te geven. Daarbij een rocknummer in en de twee systemen gaven elk iets heel anders terug. Ik probeerde dat te verklaren en te begrijpen. Daarom moet de mens dat doen. Wat willen we er dan uit komen, is dat een goede test?“

„Als valkuil kan je dit vergelijken met een spiegel in de tegenoverstaande muur. Mijn promovendus zijn er natuurlijk niet alleen maar er is ook een psychologische aspect. Daarom moet ik creatief zijn om dat te verklaren. Klaar maken dat wel verschillend. Daarom moet ik dat perspectief. Dat is echte de richting die het op moet, ook al kost het meer moeite.“

„De komende vijf jaar moet de Jonge Academie hoop houden hier veel meer te kunnen doen. Daarom moet de universiteit werken aan dat het goed uitgelopen, ik wil meer mensen bereiken en dat moet ook gaan over de verschillende disciplines. Beter dat raakt alle disciplines, zowel binnen als buiten de universiteit.“

<https://www.trouw.nl/politiek/hoe-de-belastingdienst-lage-inkomens-profileerde-in-de-jacht-op-fraude~bbb66add/>

deVerdieping Trouw

TROUW DINSDAG 23 NOVEMBER 2021

Het is de hoogste tijd dat al die onmisbare werkneemers een eerlijk loon krijgen

Europarlementair Agnes Jongerius,
economie 14

WWW.TROUW.NL | 95STE JAARGANG NR. 23639 | E-MAIL VIA WWW.TROUW.NL/SERVICE | POSTBUS: 859 1000 AW AMSTERDAM | REDITACIE: 020-562 9444 | KLANTENSERVICE: 088-056 1588 | BEZORGING: 088-056 599

Belastingdienst ging vooral achter lage inkomen aan

Fraudejacht • Om toslagen te controleren op fouten en fraude gebruikte de Belastingdienst een zelflerend algoritme. Dat selecteerde vooral mensen met lage inkomens voor controle.

Jan Kremersbusch
redacteur economie

De Belastingdienst heeft lasteling specifiek hangers met een lage inkomen geselecteerd voor extra controles op fraude met de kinderopvang-toeslag. Hogere inkomen leidt juist bewijst ontzien. Dat erkent de Belastingdienst nu toe. De dienst vroeg van Trouw en RTL Nieuws over het zogenoeten risicoclassificatiemodel.

Het risicoclassificatiemodel werd vanaf april 2013 gebruikt om te bepalen welke mensen de Belastingdienst extra controles zouden krijgen. In half voorjaar werd het uit de hoge risicoscore gehaald. Een rapport van consultancybureau KPMG, het grootste accountancybedrijf ter wereld, dat zelfstandig bepaalde welke indicatoren erop zouden kunnen wijzen dat een persoon meer kans had toslagen een verhoogde kans op fouten of fraude kennen. Alle aangevraagde toeslagen werden daarbij ingezet als een hogere risicoscore toetsbaarheid. De Belastingdienst zag de hogere risicoscore niet alleen volgens extra gecontroleerd.

De Belastingdienst stelt dat in deze periode maar 20 procent het inkomen van mensen "met variabel geschatte inkomen" had. Daarbij was praktijk berken dat die lage inkomen een hogere risicoscore kregen, hoewel de Belastingdienst een hogere lage risicoscore. Volgens de Belastingdienst bleek er "een statistisch verband te bestaan tussen de lage inkomen en de kans op een (onjuist) aanspraak".

In de loop van 2013 tot 2016 is

In het kort
De Belastingdienst gebruikte bij de controles op fraude een zelflerend algoritme

In de praktijk was dat model liegarme inkomen aan als risicotraject

Er was geen proces om deze selectiebasis tegen te gaan

Als de uitkomst zo inzoomt op een specifiek groep dan moeten er alarmbellen afgaan

Cynthia Liem
universitair hoofddocent
kognitieve intelligentie

82,3

procent van de groep met een hoge risicoscore had een inkomen van minder dan 20.000 euro per jaar. Dat is ruim 10 procent meer dan de rest van alle aangemaakte had slechts 7,3 procent een huishoudinkomen van minder dan 20.000 euro per jaar.

Selectiebasis

In de praktijk bleek het model toch voor laagere inkomen de hoogste risicoscore te hebben. Dat betekent dat commentatoren die dat vrijgaven de Belastingdienst juist bewijst ontzien. De huishoudinkomen van minder dan 20.000 euro per jaar. Dat is ruim 10 procent meer dan de rest van

Experten op het gebied van verantwoorde toepassing van zelflerende algoritmes reageren onzett op die uitkomst. "Hoe het model precies werkt is voor ons als buitenstaander niet informatie na te gaan. Maar als de uitkomst zo inzoomt op een specifieke groep als dat zijn inkomen, dan moeten er alarmbellen gaan. Want dat betekent dat die modellen werken", reageert Cynthia Liem, universitair hoofddocent en hoofd van de kennisgroep kognitieve intelligentie aan de Technische Universiteit Delft.

Volgens Liem wijst de enorme oververtegenwoordiging van lage inkomen op gezochtere selectiebasis in het model. "Hetzelfde model verwacht dat bij hogere inkomen mensen kan is om te wonen in een goed huis en goed gezorgd te worden. Dat leidt ertoe dat vooral aannemers en huurders worden gepercipieerd. Onder het model wordt getracht wordt niet die uitkomst te verkrijgen dat mensen in een gezond huis op lage inkomen."

Geleidelijk die iets weet van huisconditie en hoeveel inkomen men kan op selectiebasis", zegt Liem. Bijvoorbeeld door ook aankondigingen te maken dat de Belastingdienst de niet bij Toeslagen, zo blijkt uit een aantal voorbeelden, maar ook uit de maart van dat jaar dat de Belastingdienst ter inzag gaf aan Trouw en RTL dat de Belastingdienst alleen de hoge scores en er is geen toetsprocedures dat ook andere kenners en marktexperts "schrikken", schrijft de dienst over het resultaat van de analyse van selectiebasis.

Verzoek om toezichtsradicatie

Belastingdienst wil dit onderzoek van Trouw en RTL. Nieuws dat de Belastingdienst ook een indicator gehanteerde voor de kans op fraude al dan niet hebben van de Nederlandse nationaliteit. Maar aanleiding van beschuldigingen dat de Belastingdienst daarvan door de Autoriteit Persvoorzieningen en Amsterdams Intercity (APAI) is veroordeeld. De politieke partij Recht Leven (SP) eind oktober al verzocht dat het model openbaar te maken. De politieke partijen GroenLinks en de Partij van de Arbeid (PvdA) en de PvdA's Alexander van Huffelen (toeslagadviseur) en de PvdA's Pauline Kamer gaan, zo liet zij weten.

de Verdieping 23
Hoe de Belastingdienst
lage inkomen profiteerde

Plannen voor code zwart

Ziekenhuizen in de knel
vandaag 4/5

Hoe gaat een agent om met rellen?

Geestelijke zorg bij
hulpdiensten
vandaag 7

Criminelen
werken op
feestdagen

Online oplichters
schakelen tandem bij
nederland 9

Nieuw huis voor
de Schreeuw

Olo heeft een Munchmuseum

de Verdieping 23

Ontdek Trouw in
optima forma

L. Goff 4.124
24/11/21

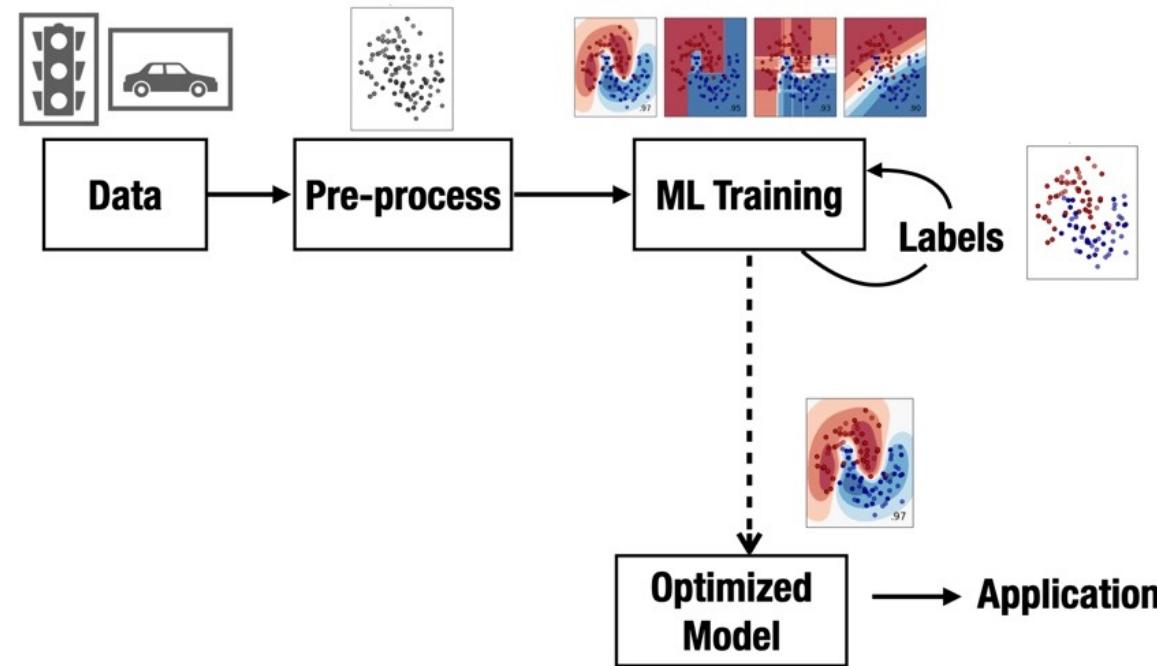
ECP

What I am encountering

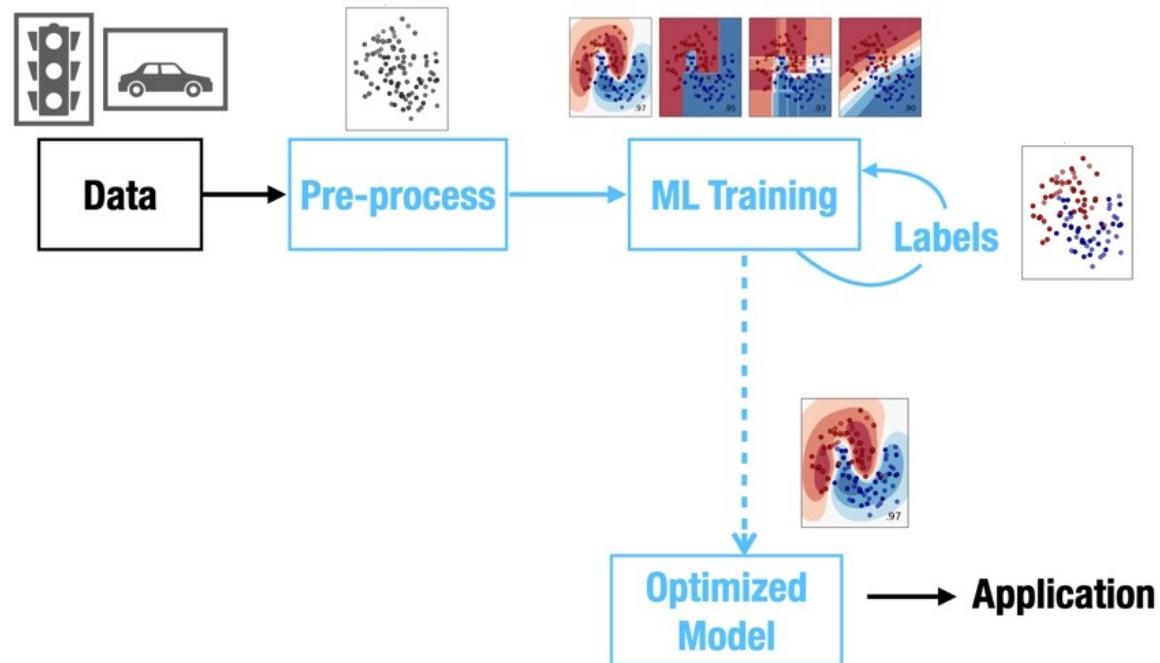


Communicating interdisciplinarily

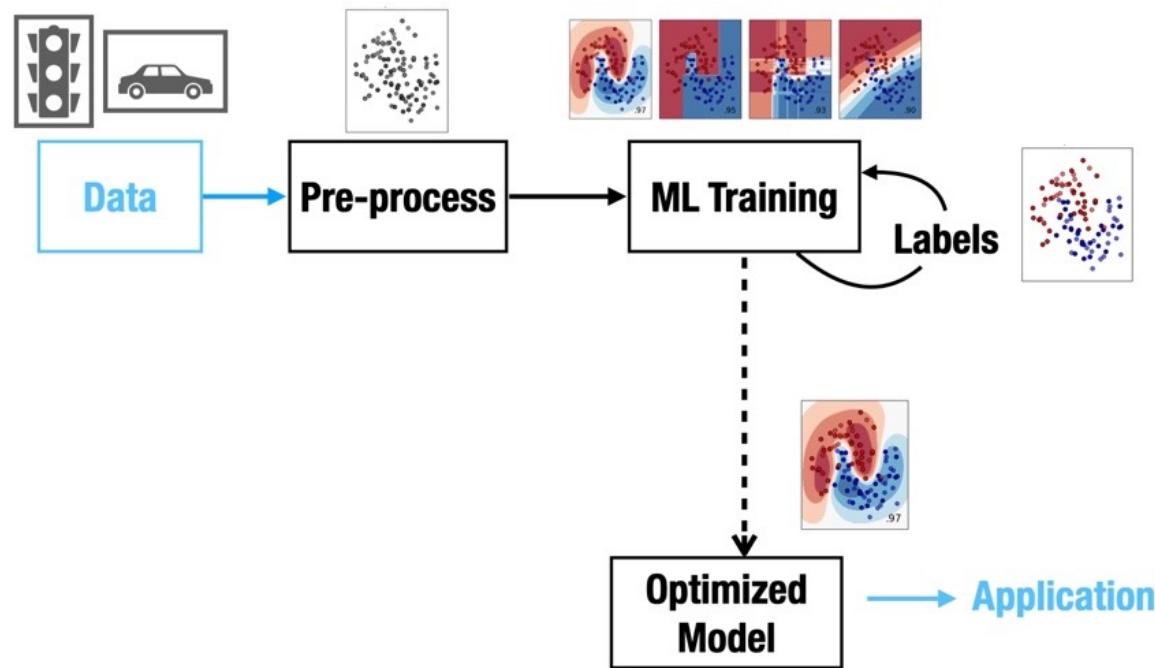
The setup



The ML expert's focus



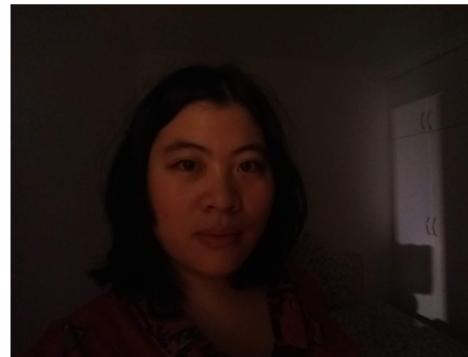
The domain expert's focus



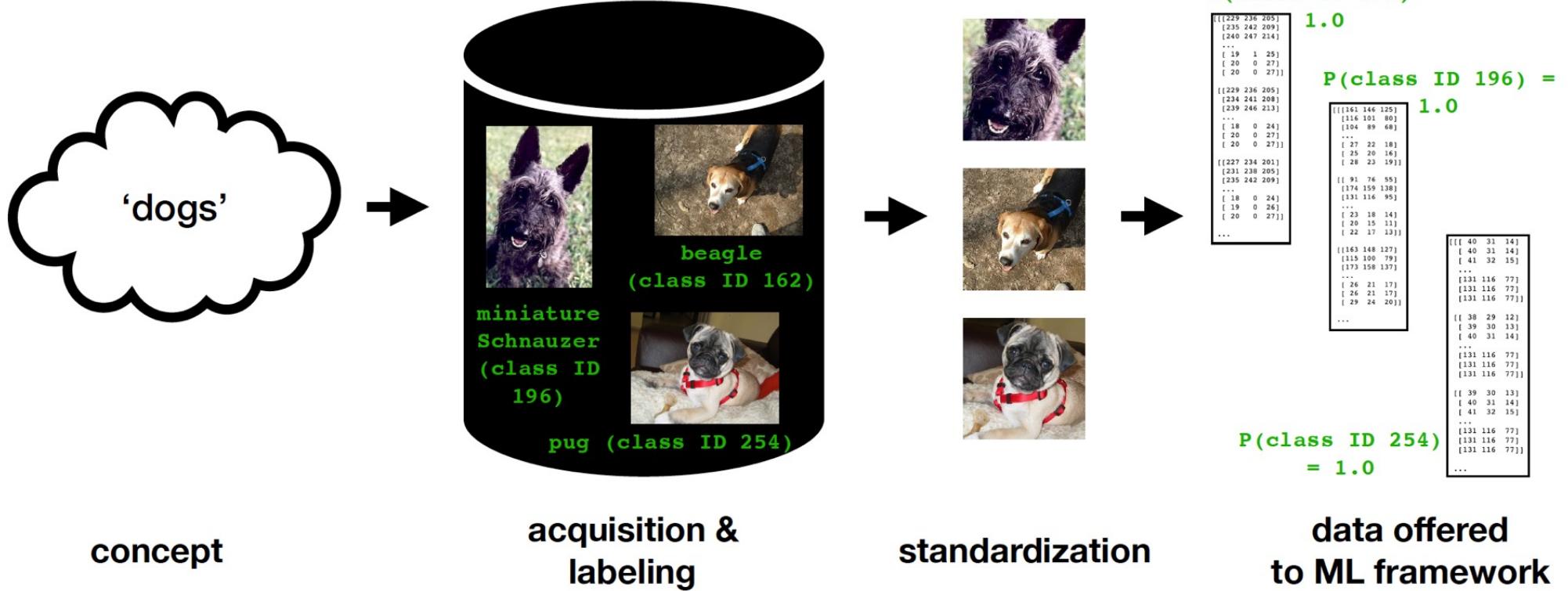
The ‘semantic gap’



The ‘semantic gap’



Standardization



Standardization

	.mp4	538.6 MB	MPEG-4 movie
	.mov	104.3 MB	QT movie
	.MOV	164.2 MB	QT movie
	.mp4	8.9 MB	MPEG-4 movie
	.mp4	7.8 MB	MPEG-4 movie
	.mp4	96.8 MB	MPEG-4 movie
	.MOV	3.7 MB	QT movie
	.mp4	19.5 MB	MPEG-4 movie
	.mp4	5.2 MB	MPEG-4 movie
	.wmv	2.9 MB	Windo...ia Video
	.mp4	105.6 MB	MPEG-4 movie
	.mov	33.7 MB	QT movie
	.MOV	7.7 MB	QT movie

Standardization

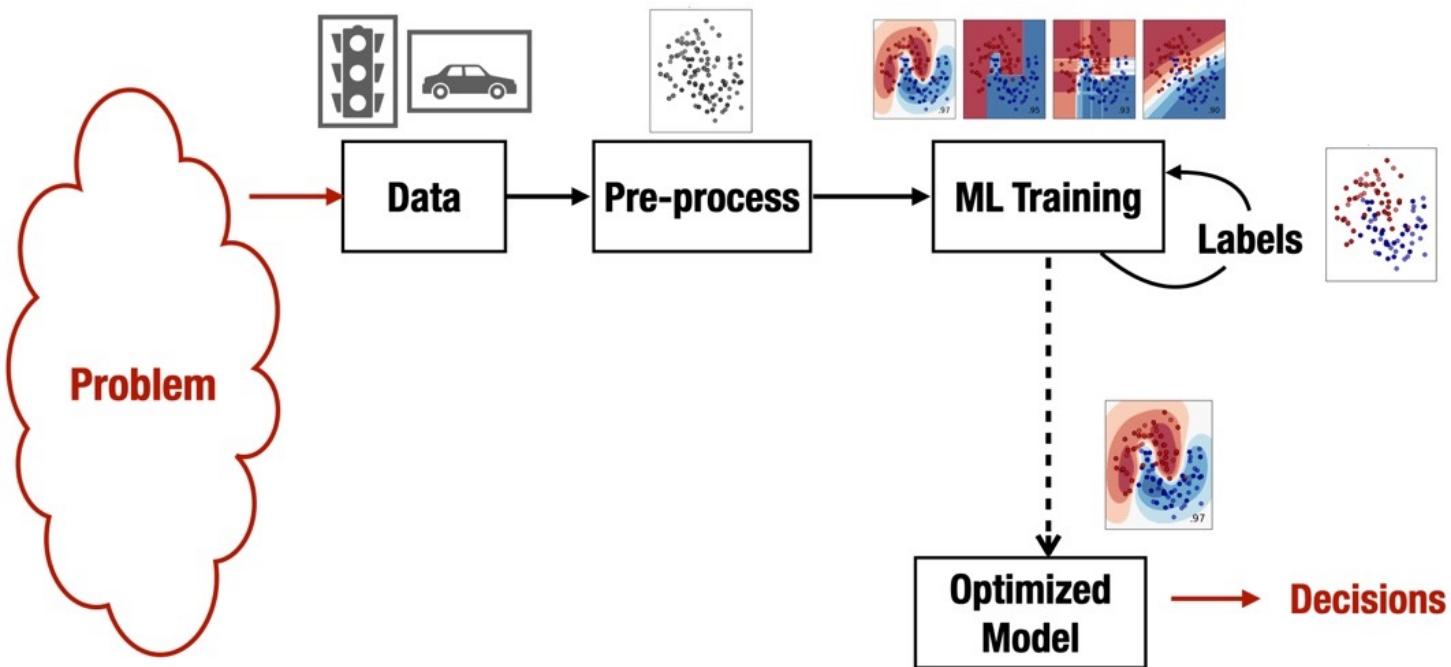


original

cropped

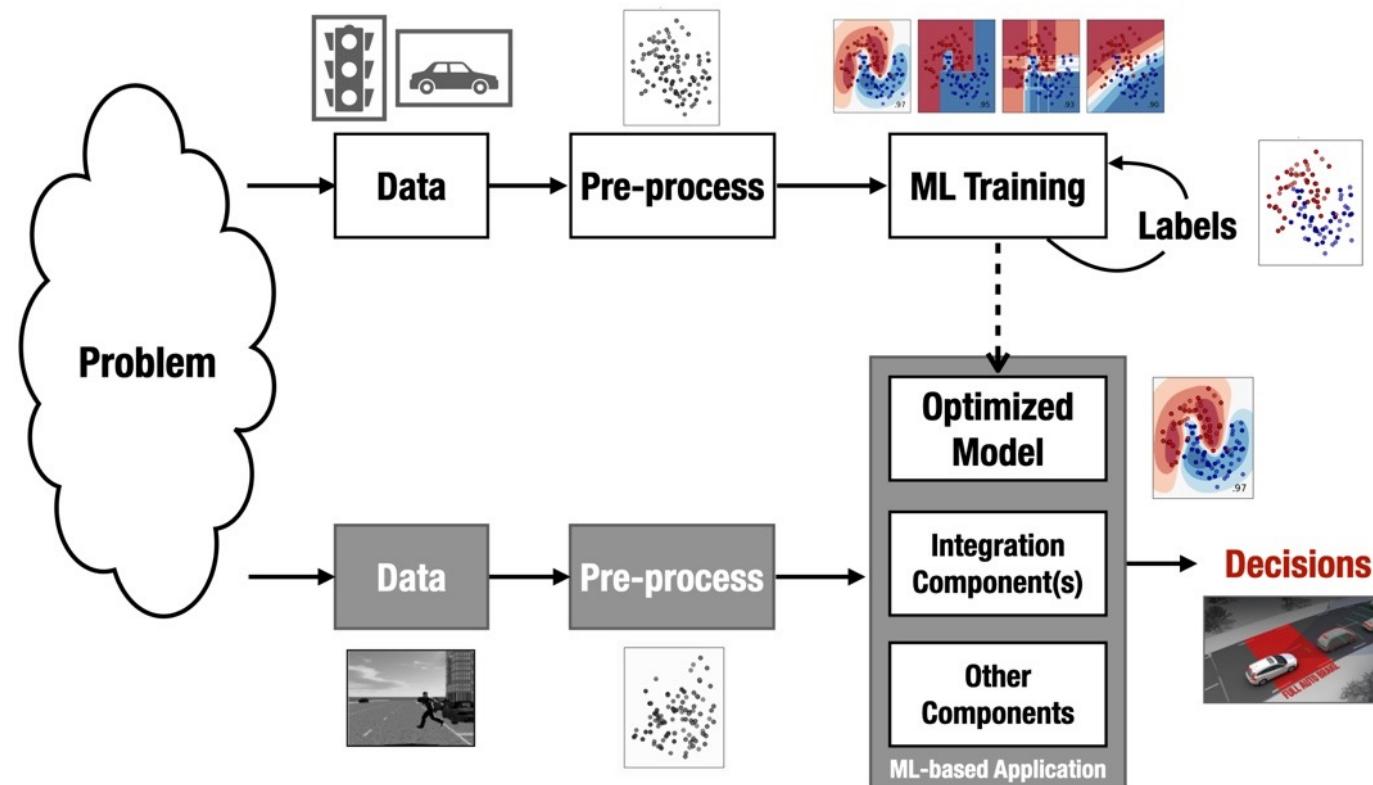
scaled

What it really may be about



What is ‘working as intended’? What is ‘the code’?

joint work on testing for ML with Annibale Panichella



Communicating for accountability

“I want to know exactly what the algorithm is doing”



UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

Source code of used software and algorithms

- ‘We use Python 3.7.5’
- ‘We also use SQL server 2016. It is licensed by Microsoft, we cannot share source code’
- ‘Python is open source so we won’t share code either’
- ‘Model can be found at https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/ensemble/hist_gradient_boosting/gradient_boosting.py’

UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

Source code of used software and algorithms

- ‘We use Python 3.7.5’
- ‘We also use SQL server 2016. It is licensed by Microsoft, we cannot share source code’
- ‘Python is open source so we won’t share code either’
- ‘Model can be found at https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/ensemble/hist_gradient_boosting/gradient_boosting.py’

UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

Source code of used software and algorithms

- ‘We use Python 3.7.5’
- ‘We also use SQL server 2016. It is licensed by Microsoft, we cannot share source code’
- ‘Python is open source so we won’t share code either’
- ‘Model can be found at https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/ensemble/hist_gradient_boosting/gradient_boosting.py’

UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

Source code of used software and algorithms

- ‘We use Python 3.7.5’
- ‘We also use SQL server 2016. It is licensed by Microsoft, we cannot share source code’
- ‘Python is open source so we won’t share code either’
- ‘Model can be found at https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/ensemble/hist_gradient_boosting/gradient_boosting.py’

UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

Source code of used software and algorithms

- ‘We use Python 3.7.5’
- ‘We also use SQL server 2016. It is licensed by Microsoft, we cannot share source code’
- ‘Python is open source so we won’t share code either’
- ‘Model can be found at https://github.com/scikit-learn/scikit-learn/blob/main/sklearn/ensemble/hist_gradient_boosting/gradient_boosting.py’

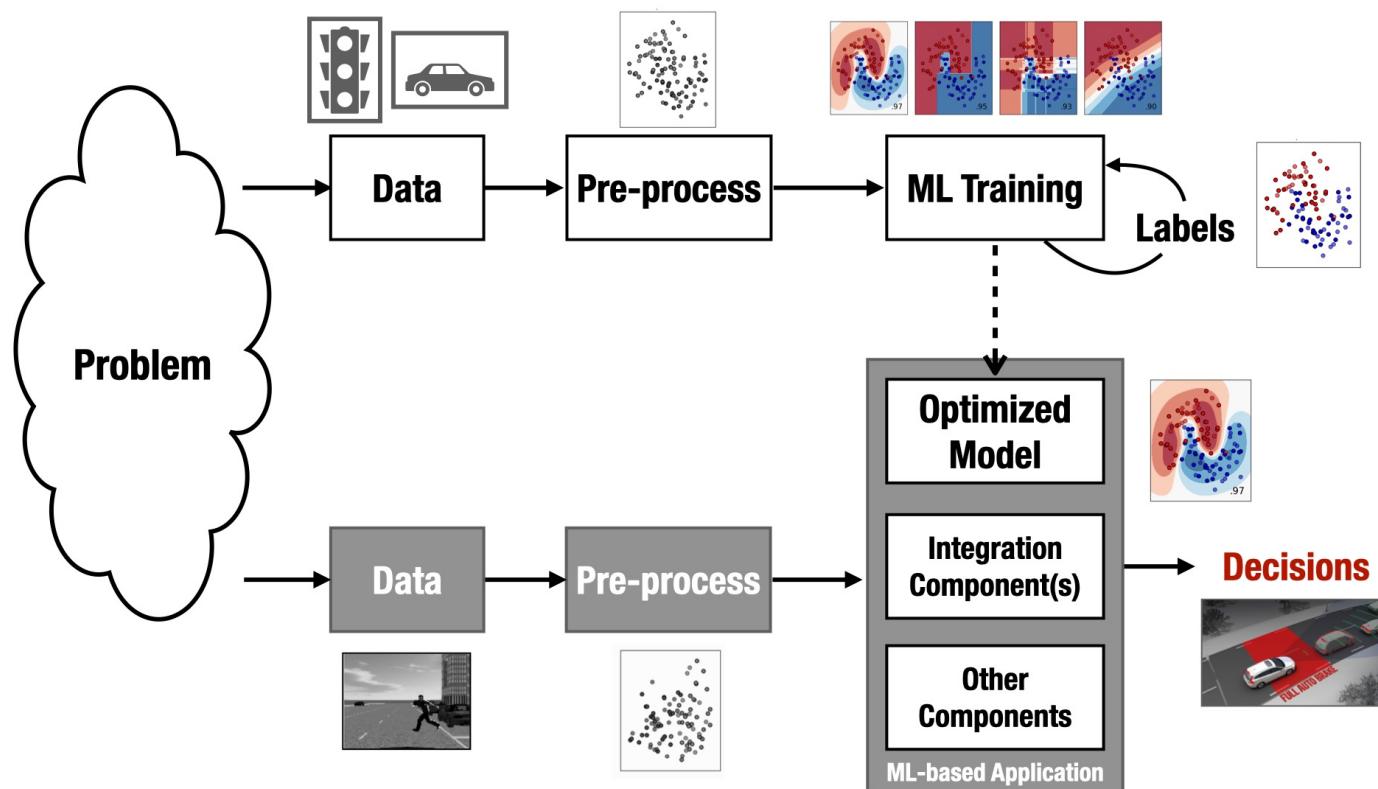
UWV (Employee Insurance Agency)

For the Dutch-reading audience: <https://www.uwv.nl/overuwv/Images/besluit-wob-verzoek-inzake-software-en-algoritmes.pdf>

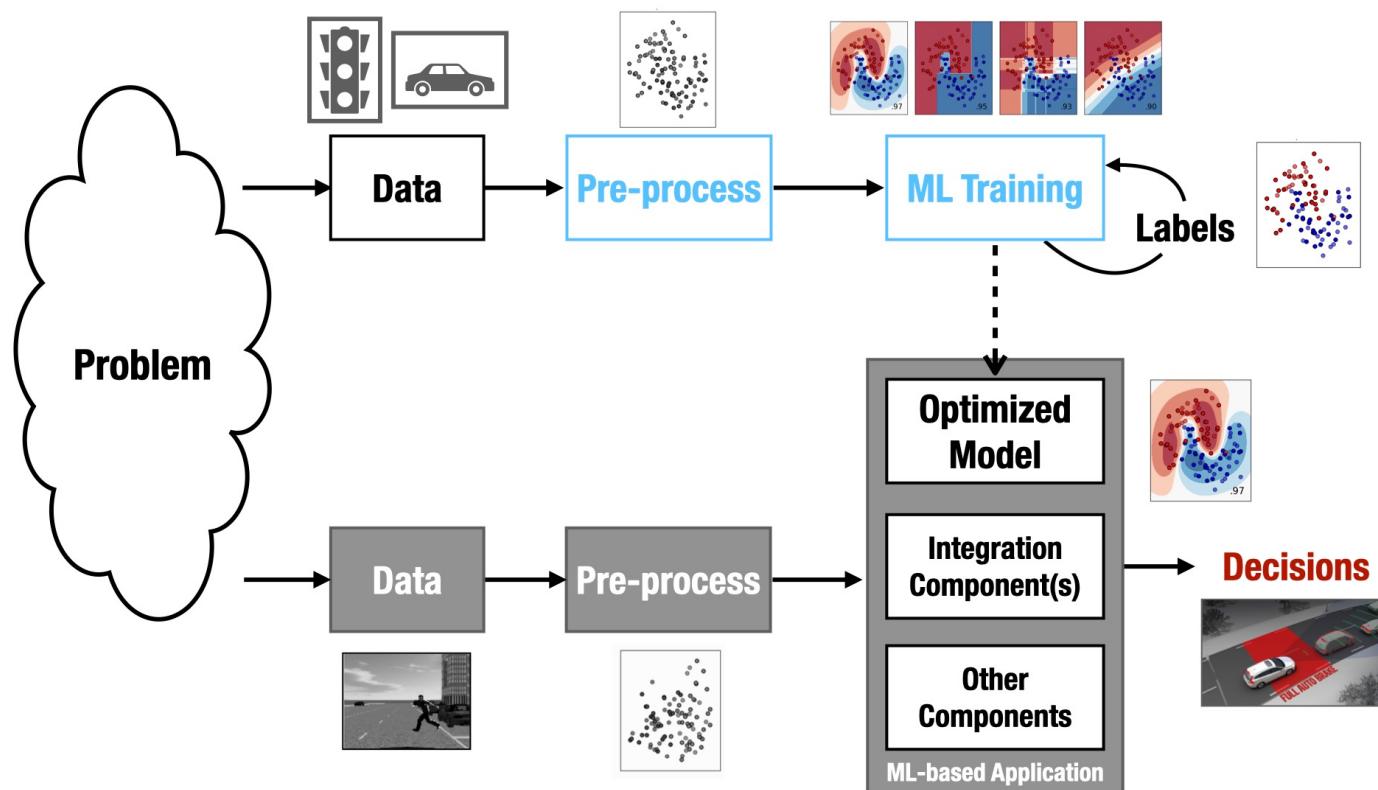
Source code of used software and algorithms

- ‘We have algorithm documentation, scripts and a Data Protection Impact Assessment. These will not be shared because of their strategic importance to our servicing.’

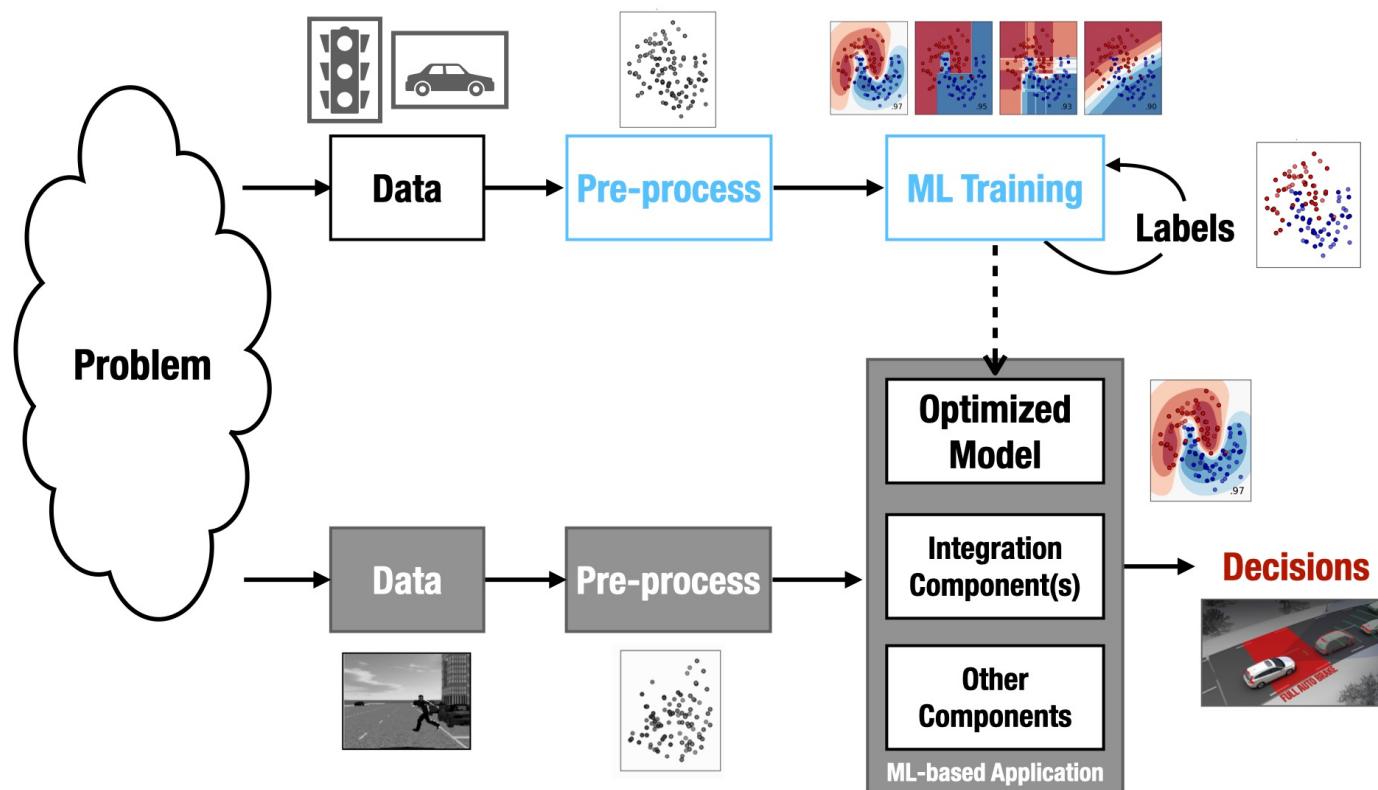
'Can you walk me through the code?'



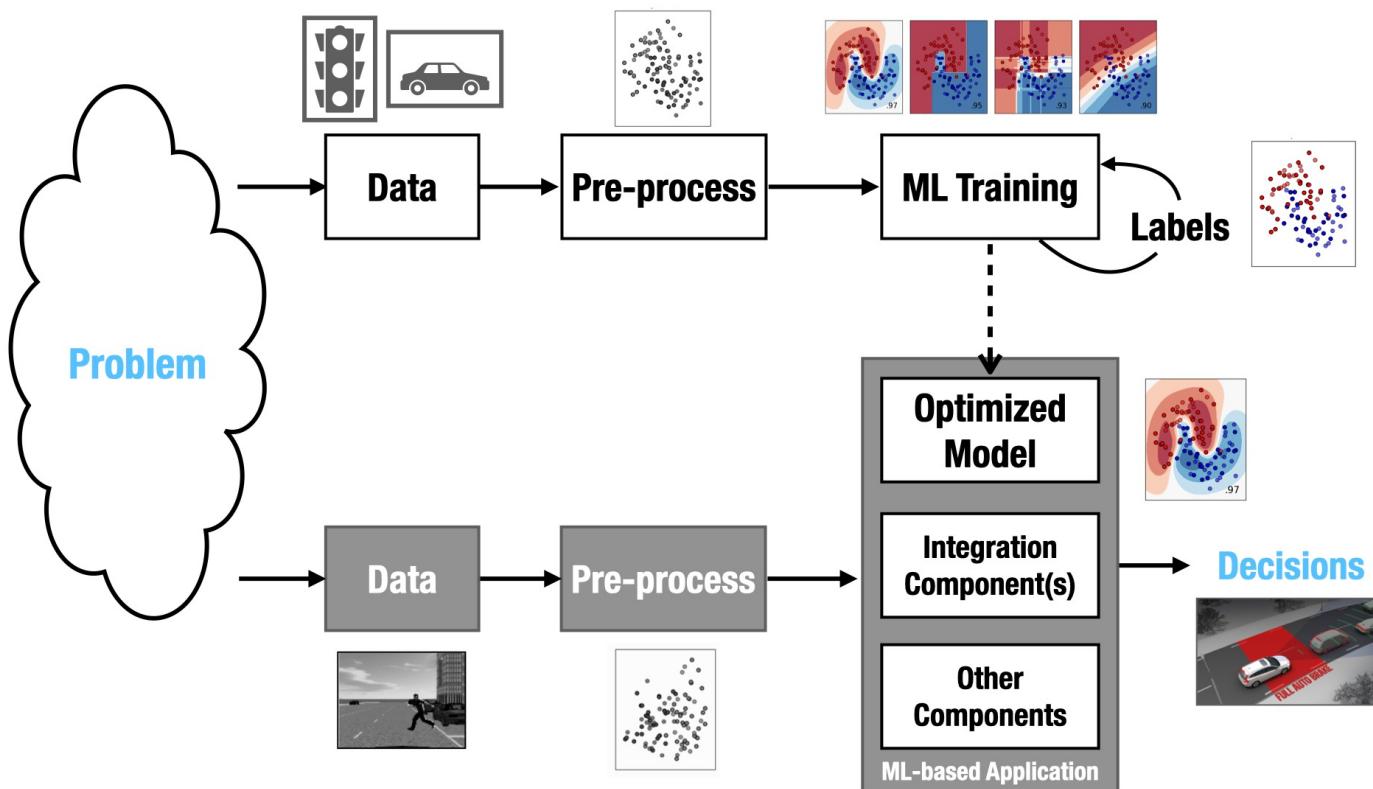
'Can you walk me through the code?'



'Should this sensitive attribute have been included?'



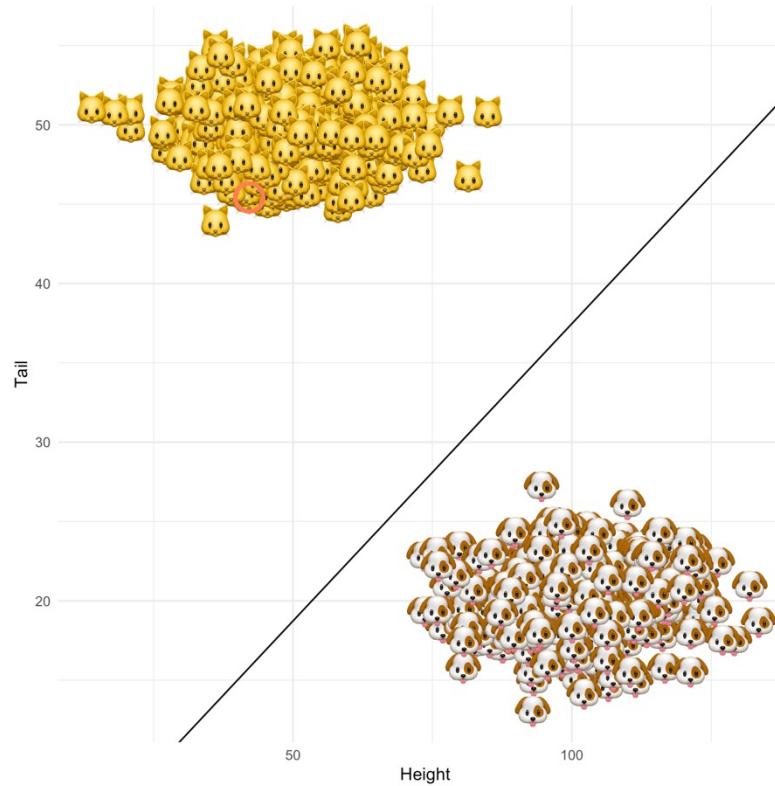
'Should this sensitive attribute have been included?'



Communicating to laypeople

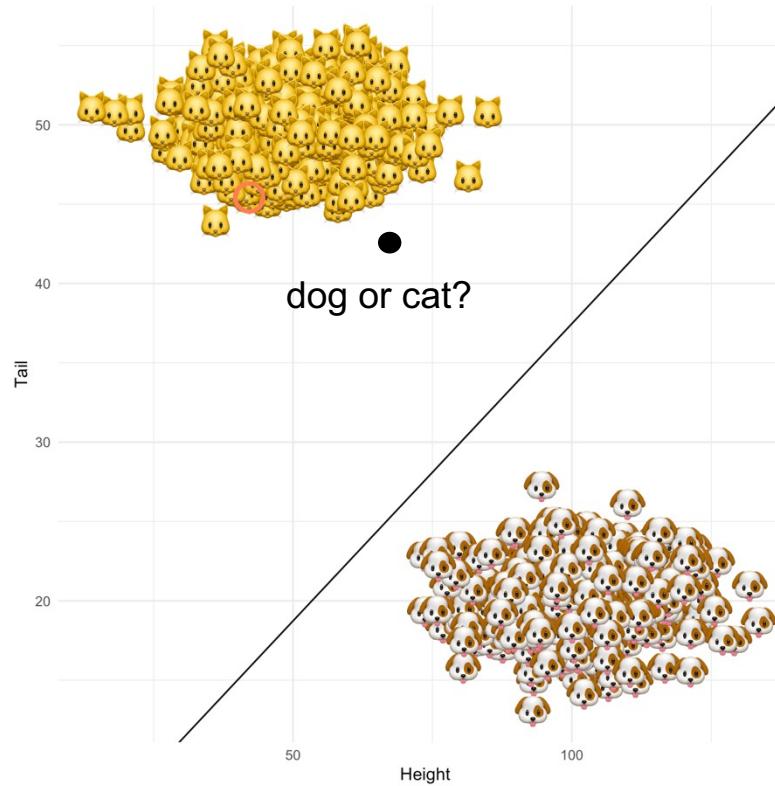
Learning patterns from data

illustration by Patrick Altmeier

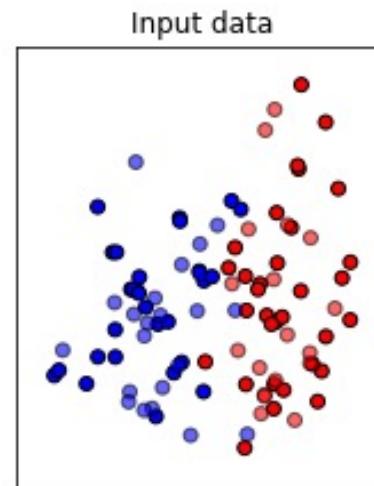


Learning patterns from data

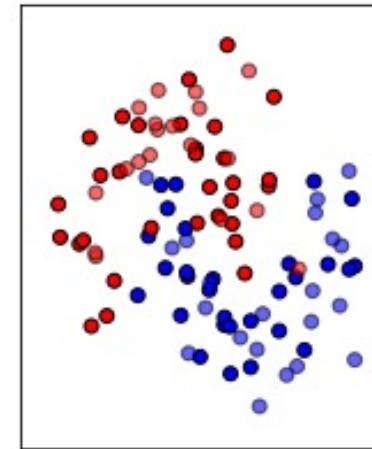
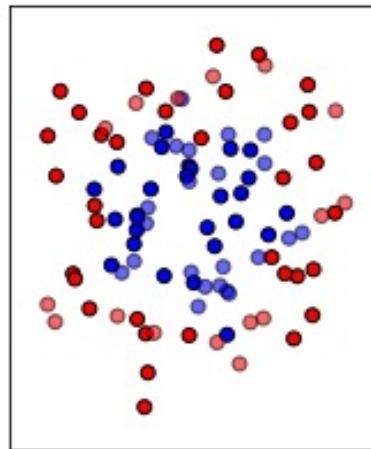
illustration by Patrick Altmeier



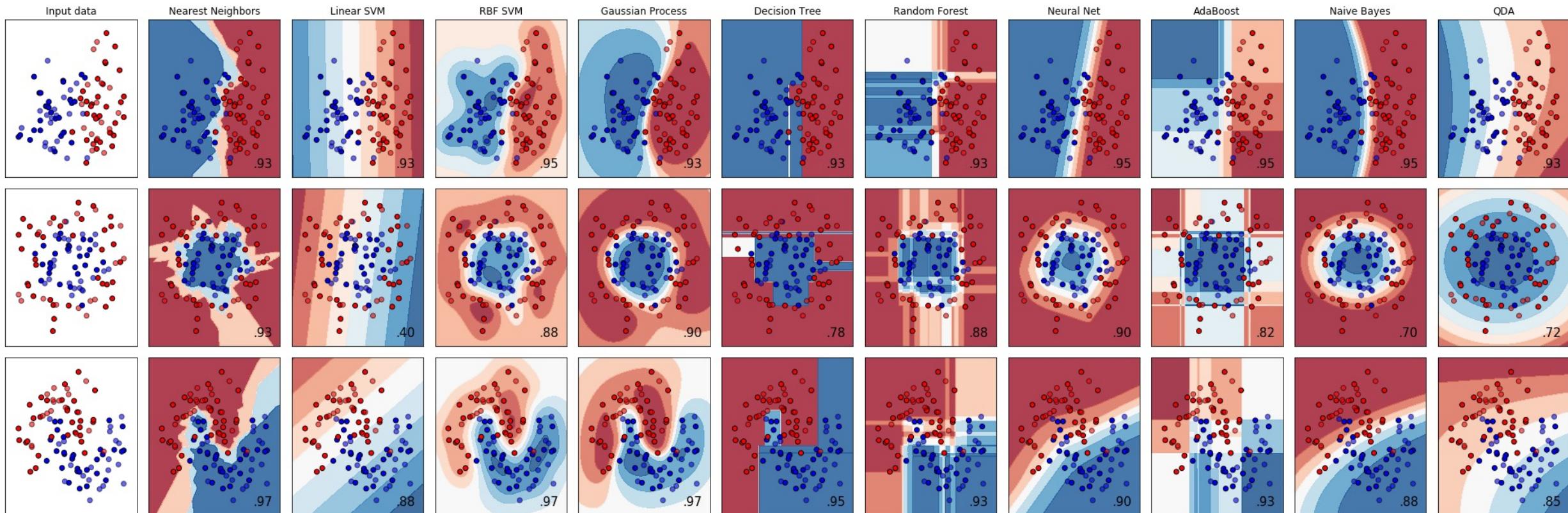
Learning patterns from data



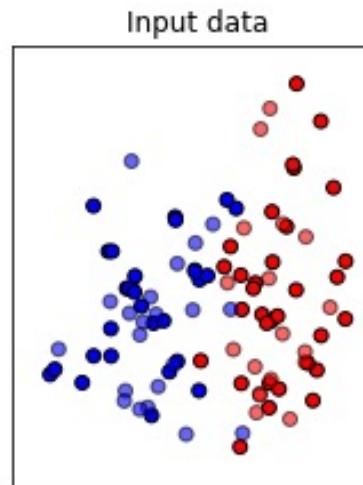
Learning patterns from data



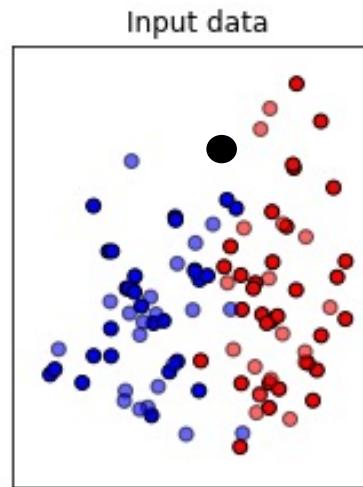
Learning patterns from data



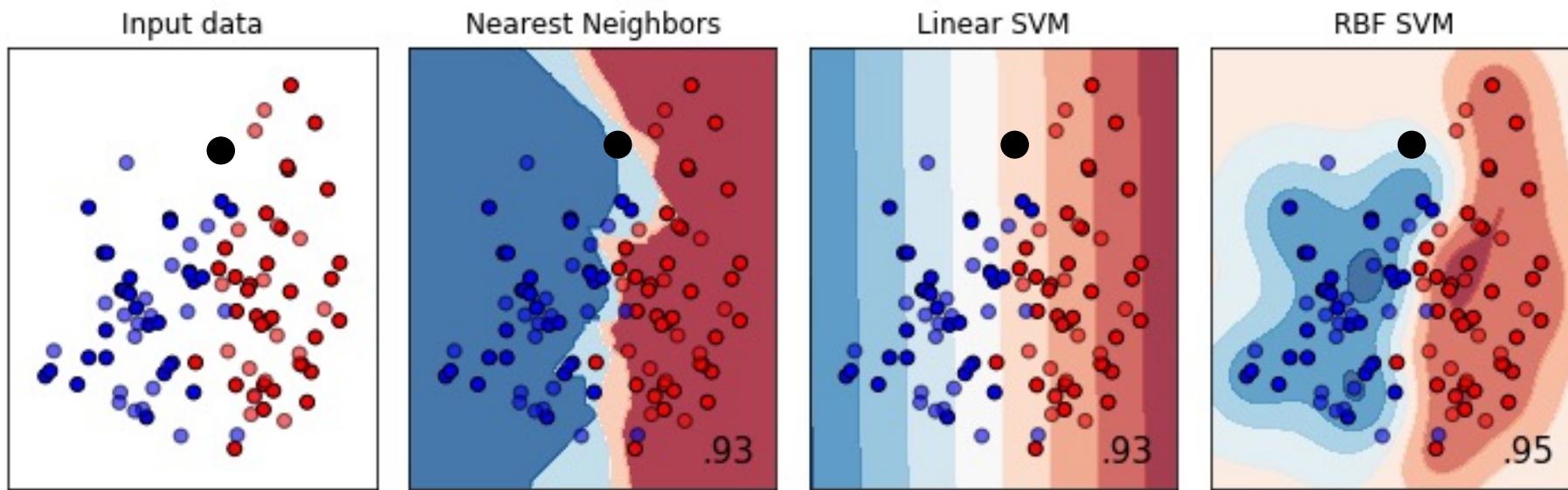
Learning patterns from data



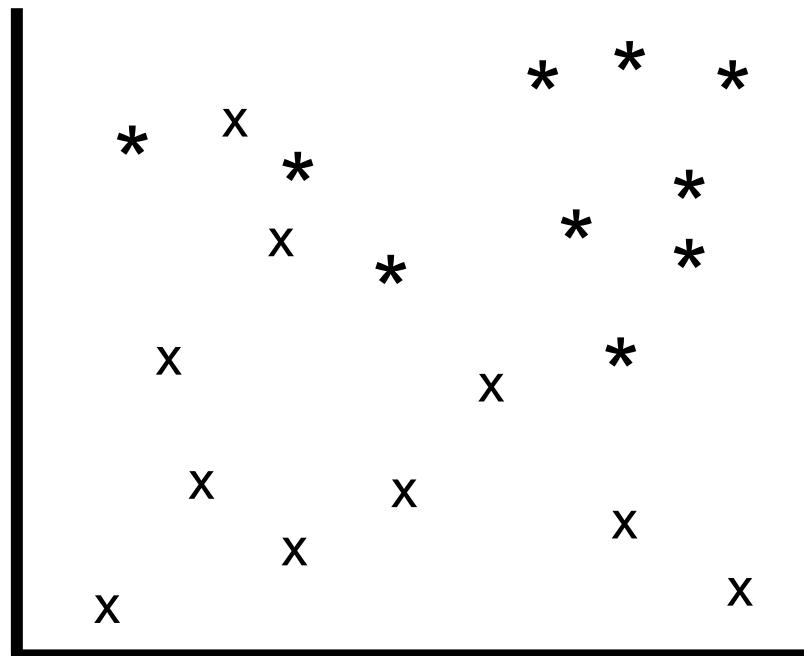
Learning patterns from data



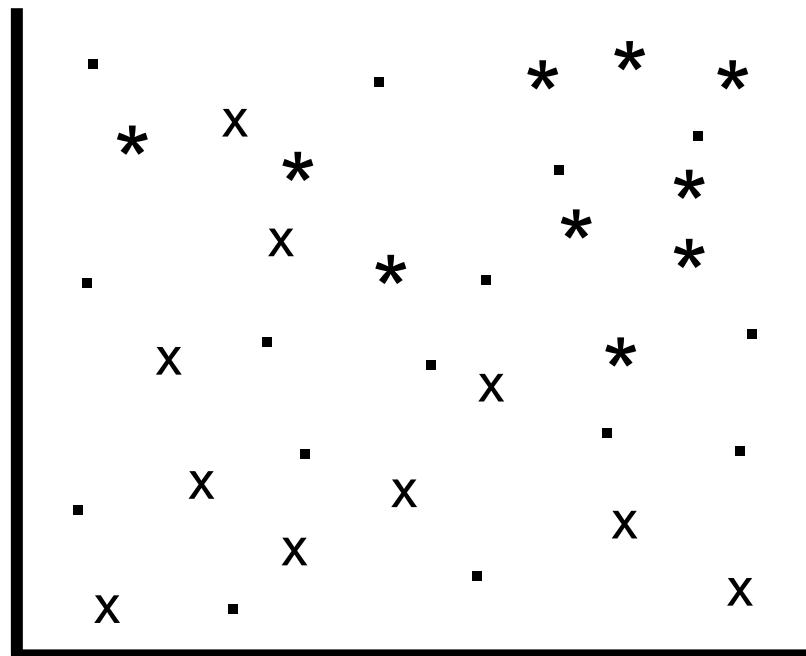
Learning patterns from data



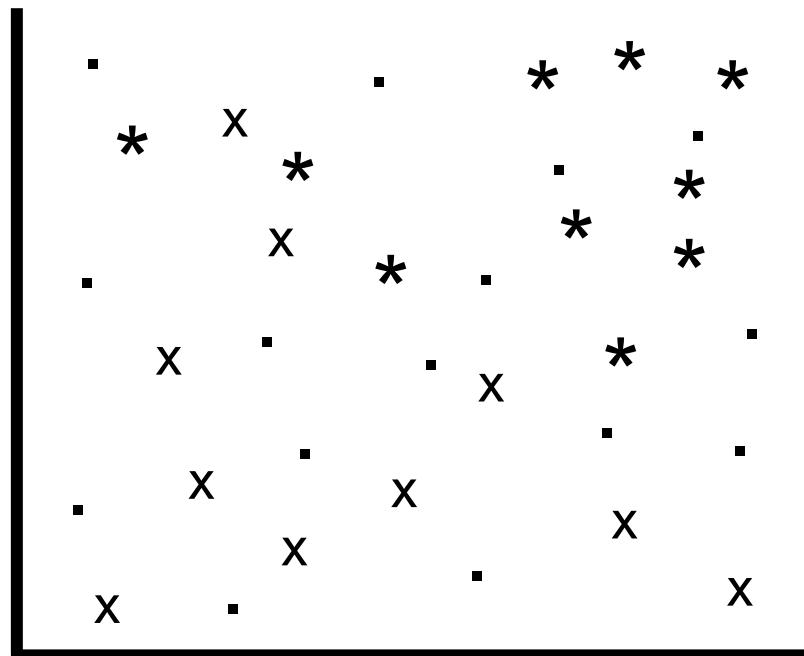
From data to decision-making



From data to decision-making



What if you only can pick 5 points?



What if you only can pick 5 points?



https://tudelft.fra1.qualtrics.com/jfe/form/SV_3I8HIOZxBVamFIW

Conclusion

We need to grow literacy.

Especially **at the interfaces** between data-driven pipelines and their surrounding contexts.

This goes beyond introductory ML and programming courses.

Conclusion





```
if questions:  
    try:  
        answer()  
    except RuntimeError:  
        pass  
else:  
    print("Thank You.")
```

References

1. C.C.S. Liem et al., Psychology meets Machine Learning: Interdisciplinary perspectives on algorithmic job candidate screening, in Explainable and Interpretable Models in Computer Vision and Machine Learning (pp. 197-253), 2018.
2. C.J. König et al., Some Advice for Psychologists Who Want to Work With Computer Scientists on Big Data, Personnel Assessment and Decisions, vol. 6 no. 1, 2020.
3. A. Panichella & C.C.S. Liem, What Are We Really Testing in Mutation Testing for Machine Learning? A Critical Reflection, 43rd Conference on International Conference on Software Engineering (ICSE) – New Ideas and Emerging Results, 2021.