



# Explainable AI

Yogesh Kulkarni  
Principal Architect – CTO Office





What we do in AIML team ...

# ICM DiscoverAI App

*AI Powered NLP & Semantics Driven Clause Discovery, Deviations and Verifications*

The screenshot displays the ICM DiscoverAI App interface. At the top, the 'icertis Applied Cloud' logo is on the left, and a search bar with 'Enter search here...' is in the center. To the right of the search bar are icons for Search, Notifications, and Help, along with a 'Welcome CLM Admin' message. Below the header is a navigation bar with icons for Dashboard, Agreements, Create Agreement, Requests, Create Request, and Bulk Upload. The main content area shows the breadcrumb 'My Dashboard > Agreement Management > Details' and the title 'Master Services Agreement - ICMMasterServicesAgreement\_104'. On the left is a sidebar with a list of categories: Agreement, Associations, All, Deviations, Amendments, Commitments, Clause Discovery (highlighted with a count of 13), Metadata Discovery, Agreement Clauses, Line Item Discovery, and Team. The main panel is titled 'Discovered Clauses Mapped' and includes filters for 'All(13)', 'Confirmed(0)', 'Review Later(0)', 'Ignore(0)', and 'Not Identified(13)', along with a 'Sort By: Label\_Sort\_Docum' dropdown. It features two columns: 'Discovered Clauses' and 'Matched Library Clauses'. The 'Discovered Clauses' column shows a clause titled 'notices,preamble' with a detailed text snippet and a 'More' link. The 'Matched Library Clauses' column shows a clause titled 'MSAS Non GMA Preamble' with a detailed text snippet and a 'Track Deviation' checkbox. At the bottom of the interface, there is a footer with 'Copyright © 2018 Icertis Inc.' and the 'icertis Contract Management' logo.

## Powerful capabilities:

- Image/PDF Conversion
- Clause and Attribute Discovery
- Accurate Clause Matching
- Deviation Tracking
- Attribute Verification
- Driven by NLP & Machine Learning

# ICM Obligation Management (AI Powered) App

*AI Powered Obligations discovery and insights into your contracts*

Agreement

Summary

Details

Preview

Versions

History

Associations

All

Resource Pricing

Obligations

Response Time

Amendments

Commitments

Clause Discovery

Metadata Discovery

**Obligation Discovery**

Agreement Clauses

Team

Discovered Obligation Mapped

Export Obligation Data Save Progress Finish

All(30) Confirmed(0) Review Later(0) Ignore(0) Not Identified(30) Visualize Agreement Category:

financial mgmt :: invoicing milestones, compliance m...

At such time, Party 1 may provide Client With an invoice for the undisputed portions of the invoice and Client shall pay such invoice upon the later of the due date of the original invoice, or Within thirty (30) days after the issuance of the revised invoice for the undisputed portions

Parent Clause Ignore

financial mgmt :: invoicing milestones, compliance mg...

Each Party's total aggregate liability for all losses, liabilities, claims and damages arising out of, or related in any way to any of the following

service delivery mgmt :: incident management, financi...

The Party receiving a notice shall attempt to cure the breach during a thirty (30) day cure period or, if the breach cannot reasonably be cured

compliance mgmt :: audits & certifications, contract a...

Client shall pay for these audits, and shall conduct them at a mutually agreed upon date within a reasonable period of time after the audit is

compliance mgmt :: audits & certifications, contract a...

Matched Library Obligations

Financial management :: Invoicing m...

In the event that no payment milestones have been agreed, Supplier shall issue an invoice at [DATE\_1] in which [PERCENT\_1] of the respective Service has been accepted. [CTS]

More Compare

Confirm Review Later

Service Delivery Management :: Personnel or staff...

Show More Panels

Copyright © 2019 Icertis Inc.

icertis Contract Management

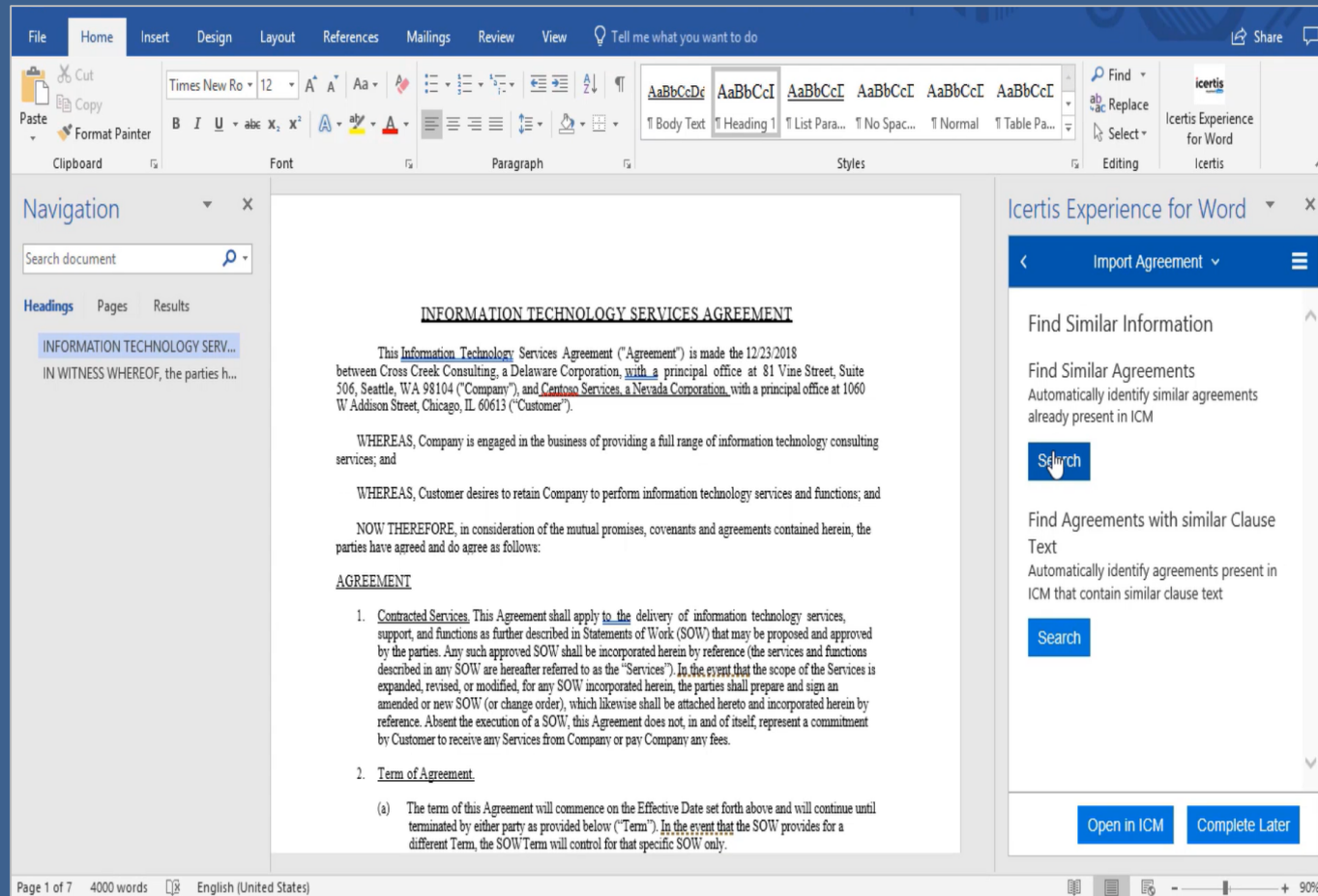
## Powerful capabilities:

- Easy Obligation Management
- Extract obligations & Categorize them with Customer Taxonomy
- Extract key meta data from discovered obligations, automatically
- Automate obligations assignment to both internal and external users
- Take direct action based on intelligent insights
- Reduced risks by complying with contractual obligations



# ICM NegotiateAI App

*AI-powered insights to optimize contract negotiation, reduce risk, and negotiate better outcomes*



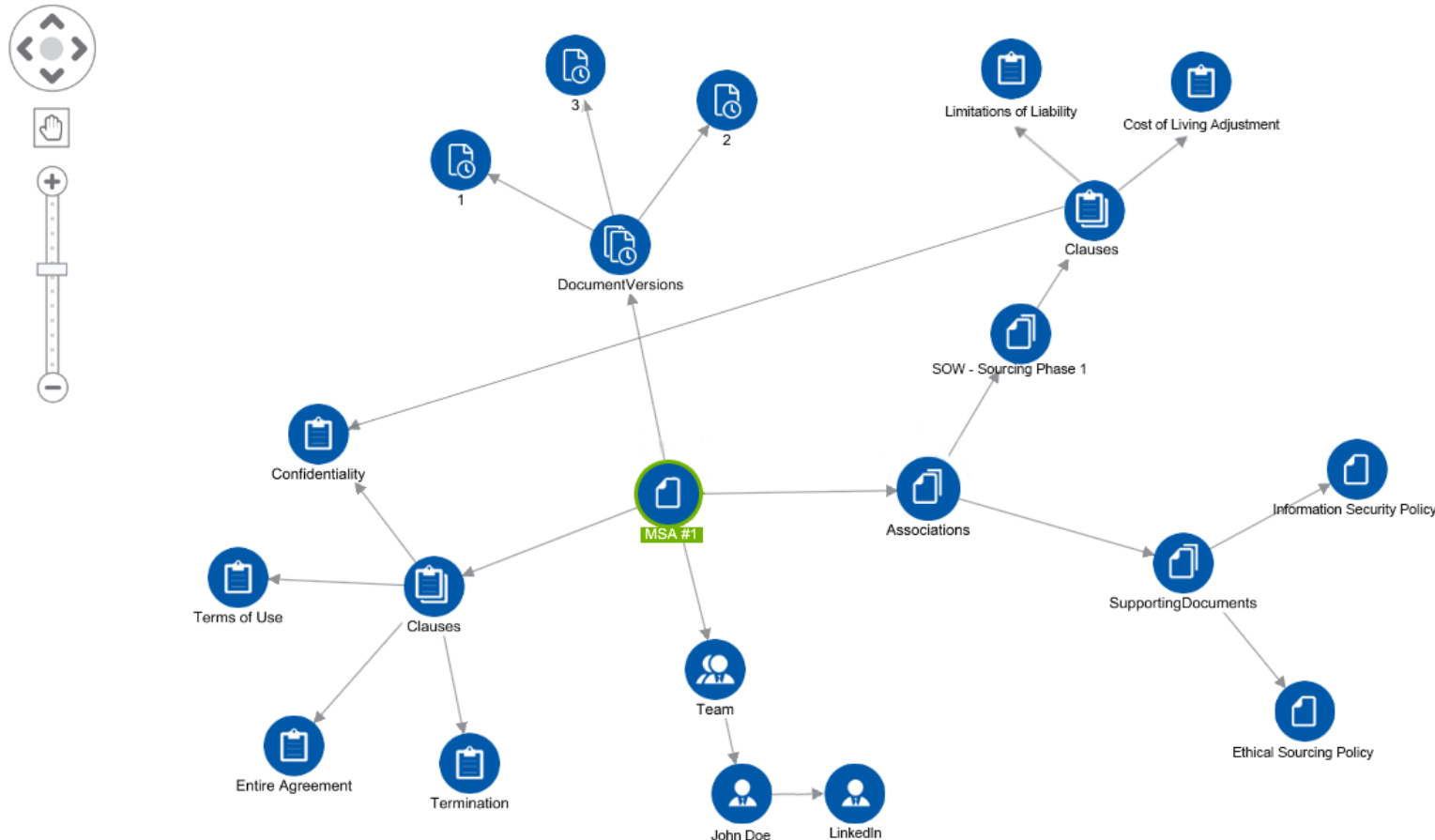
## Powerful capabilities:

- Automated contract discovery
- NLP & Semantics powered extractions
- Finds similar contract and clauses
- Highlight clause deviations to ensure compliance
- Powerful AI-infused negotiations
- Familiar Microsoft Experiences
- Efficient Business Outcomes

# ICM VisualizeAI App

*Visualize complex contract structures, relationships and connections*

icertis Visualize AI



## Powerful capabilities:

- Unified view of Contracting Across the Organization
- Auto discover dependencies, structures associations & relationships
- Unlock hidden contract insights, connections and get contextual views
- Link to external data sources
- Unprecedented insights into your entire contract portfolio
- Take direct actions based on such intelligent insights

# Overall Data Flow of Discover/Digitize AI



# A Sample Digitization Project ...

Number of Contracts to be Digitized : **60K**

Number of Clauses to be discovered : **32** (“Term”, “Warranty”, ...)

Number of Attributes to be discovered: **12** (“Effective Date”, “Contract Value”,...)

Time : **3 months**



# Steps ...

Build AI Engine using annotated samples to extract attributes

Show results of training to the Customer and get approval

Start extractions (production) for all the contracts, in batches ...

Halfway through the production ...

# A sample dialog

Customer: The extractions are looking ok, but...

AIML team: But??

Customer: Why are you digitizing these 'DELLA' contracts?

AIML team: Meaning?

Customer: You should not digitize contracts having 'DELLA Corporation' in the footer.

AIML team: But you never told us. No mention of any such rules in SOW also..

Customer: We are just asking you to NOT process these, not adding to your work/scope!!

AIML team: But....

Customer: Shouldn't be difficult, right? ***AI will find out...***

(we keep hearing...)

AI will find out!!



(sometimes...)

AI should find out!!

(what if...)

AI can not find out!!

(and worse if ...)

AI finds it and finds it wrongly!!

(bottom line... we need to understand that)

**AI is not Magic!!**



(But still, you need to explain me)

# Why AI gave this result!!

(that's)

# Explainable AI

# Need for Explainable AI



Explainable AI is essential for customers to understand and trust the decisions by AI.

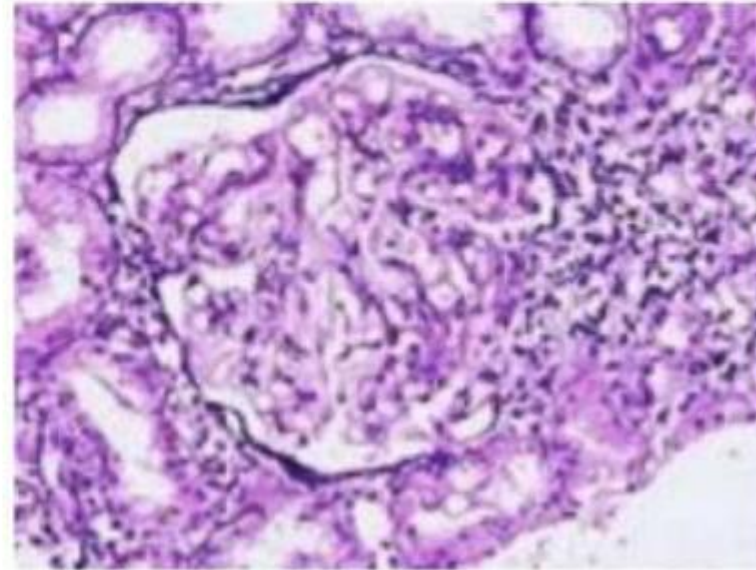
(Explainable AI in Industry, KDD 2019 Tutorial, Sahin Cem Geyik, Krishnaram Kenthapadi & Varun Mithal )

# Wrong decisions: Costly and dangerous

*“Autonomous car crashes,  
because it wrongly recognizes ...”*



*“AI medical diagnosis system  
misclassifies patient’s disease ...”*



(Credit: Samek, Binder, Tutorial on Interpretable ML, MICCAI'18)



# AI missed!!

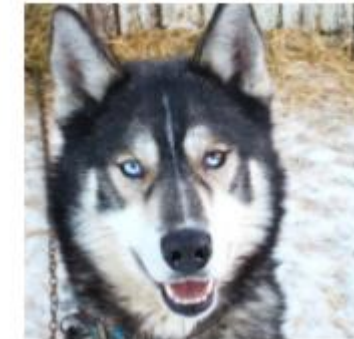
## Dog vs Wolf classifier

Instead of animal body features it started detecting based on 'presence of snow'.

So Siberian Husky (a dog) got classified as Wolf.

Training data of wolves was with snow landscapes.

<http://innovation.uci.edu/2017/08/husky-or-wolf-using-a-black-box-learning-model-to-avoid-adoption-errors/>



(a) Husky classified as wolf



(b) Explanation

**Figure 11: Raw data and explanation of a bad model's prediction in the "Husky vs Wolf" task.**

	Before	After
Trusted the bad model	10 out of 27	3 out of 27
Snow as a potential feature	12 out of 27	25 out of 27

**Table 2: "Husky vs Wolf" experiment results.**

-- “AI can make decisions better, but  
Can it make better decisions”--

# AI as Black Box

Why did the AI system do that?

Why didn't the AI system do something else?

When did the AI system succeed?

When did the AI system fail?

When does the AI system give enough confidence in the decision that you can trust it?

How can the AI system correct an error?

Explainable AI (XAI) – A Perspective, Saurabh Kaushik



# Explainable AI



# Three Key Pillars of Explainable AI

Reasonable AI: The ability to understand the reasoning behind each individual prediction

Traceable AI: The ability to trace prediction process from algorithm to data.

Understandable AI: The ability to fully understand the AI decision-making is based

“Right to Explanation” – GDPR

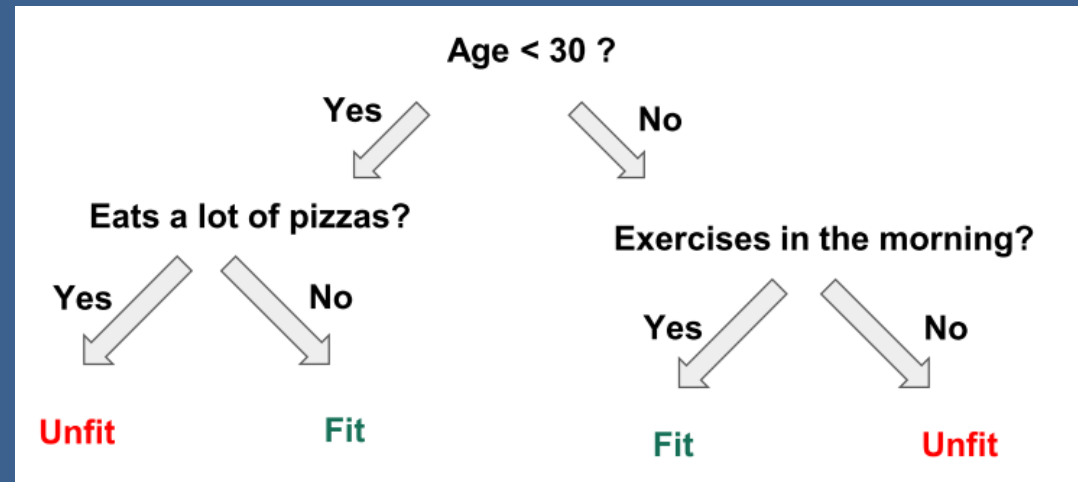
Explainable AI (XAI) – A Perspective, Saurabh Kaushik

# Achieving Explainable AI

## Prediction explanations

- Important features
- Features weights
- Decision tree plotting

Build an interpretable model using libraries like LIME, SHAP, etc.



Explainable AI in Industry, KDD 2019 Tutorial, Sahin Cem Geyik, Krishnaram Kenthapadi & Varun Mithal

**“Interpretability is the degree to which a  
human can  
understand the cause of a decision.”**

(Miller, Tim. 2017. “Explanation in Artificial Intelligence: Insights from the Social Sciences.”)



So, What (exactly) is AI?

(typical understanding)

# If Machines show intelligence, like Humans that's AI



## ARTIFICIAL INTELLIGENCE

Early artificial intelligence stirs excitement.



## MACHINE LEARNING

Machine learning begins to flourish.



## DEEP LEARNING

Deep learning breakthroughs drive AI boom.



1950's

1960's

1970's

1980's

1990's

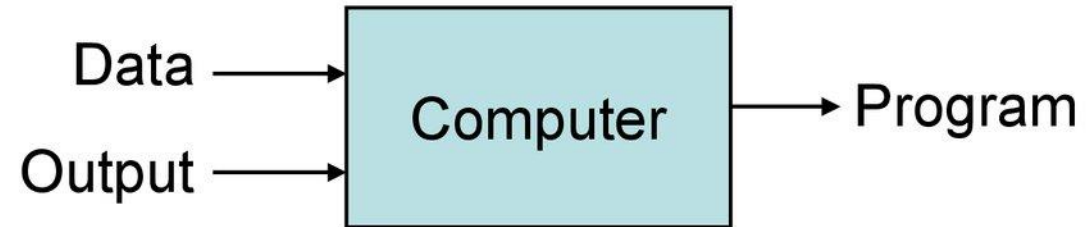
2000's

2010's

## Traditional Programming



## Machine Learning








Test



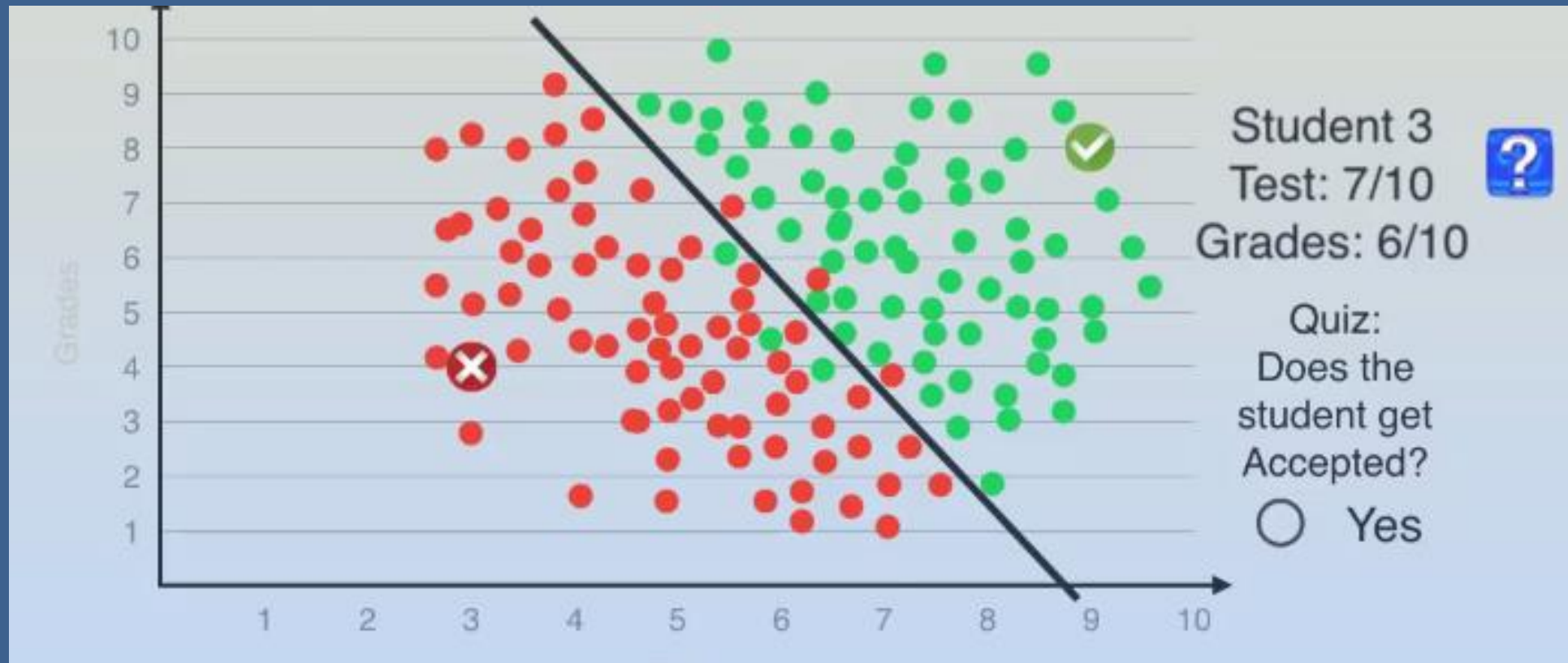
Grades

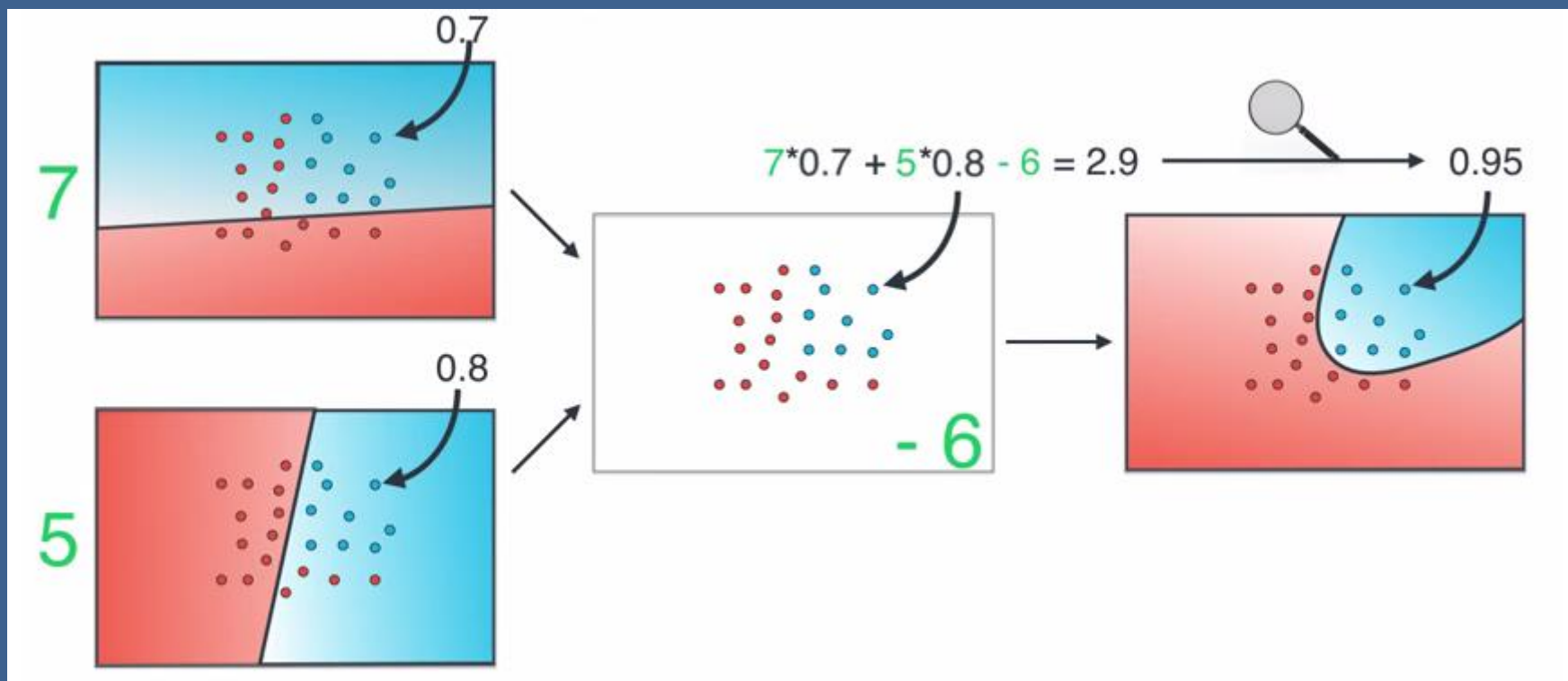
Student 1  
Test: 9/10   
Grades: 8/10

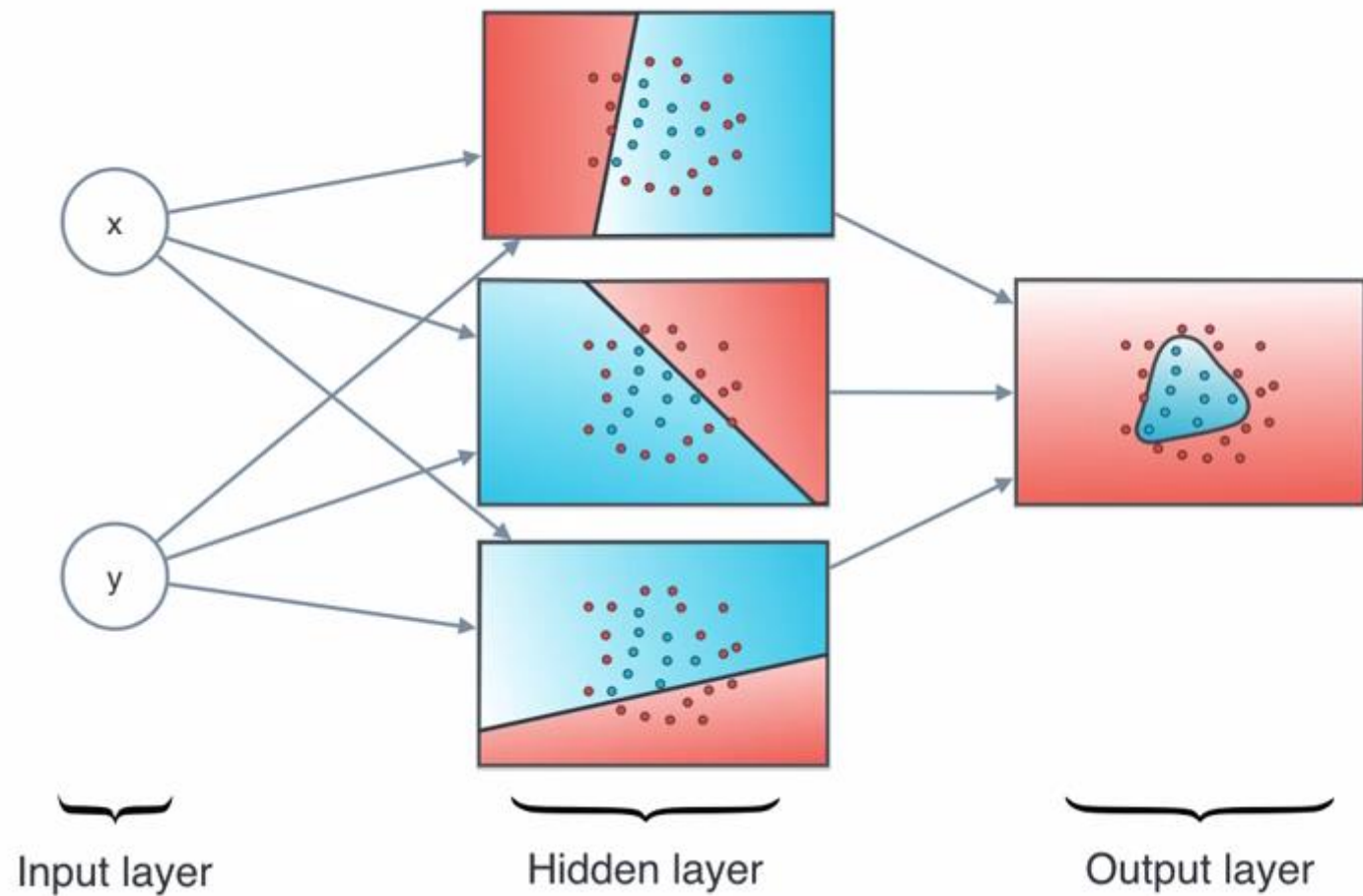
Student 2  
Test: 3/10   
Grades: 4/10

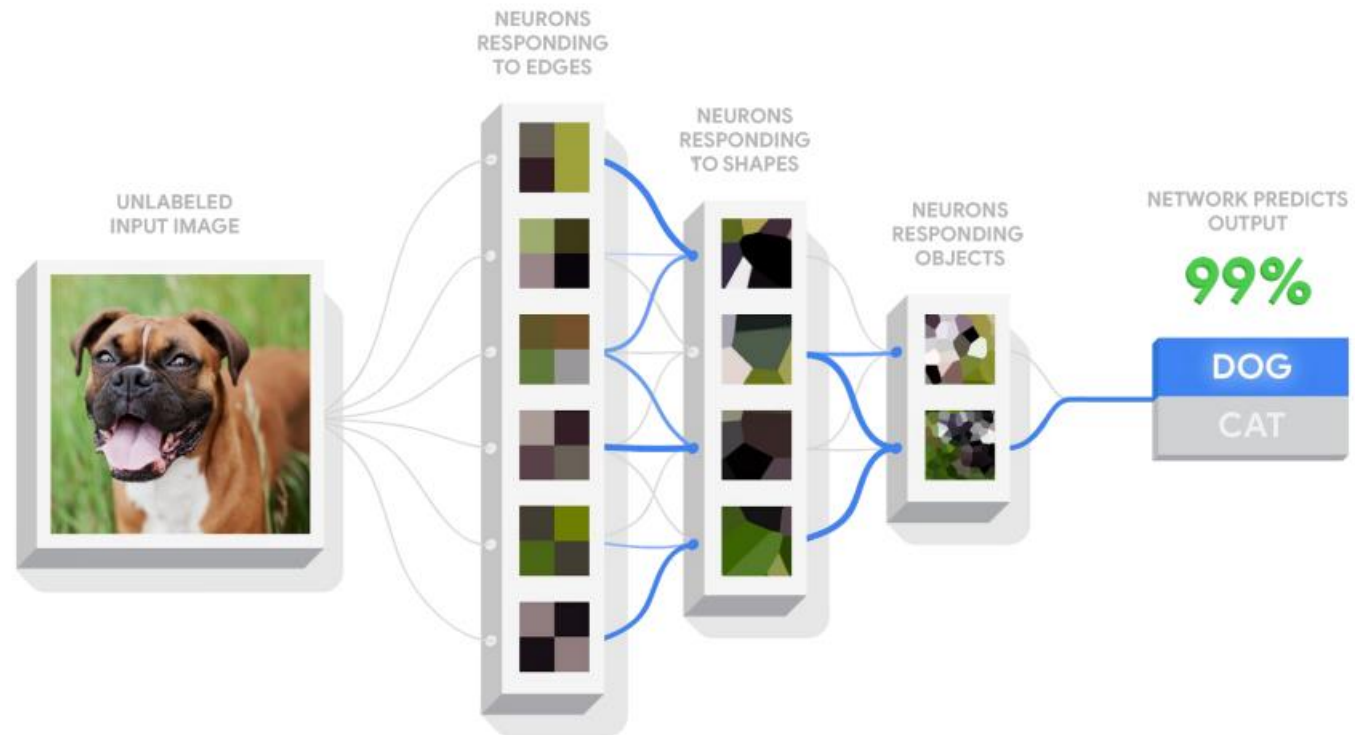
Student 3  
Test: 7/10   
Grades: 6/10



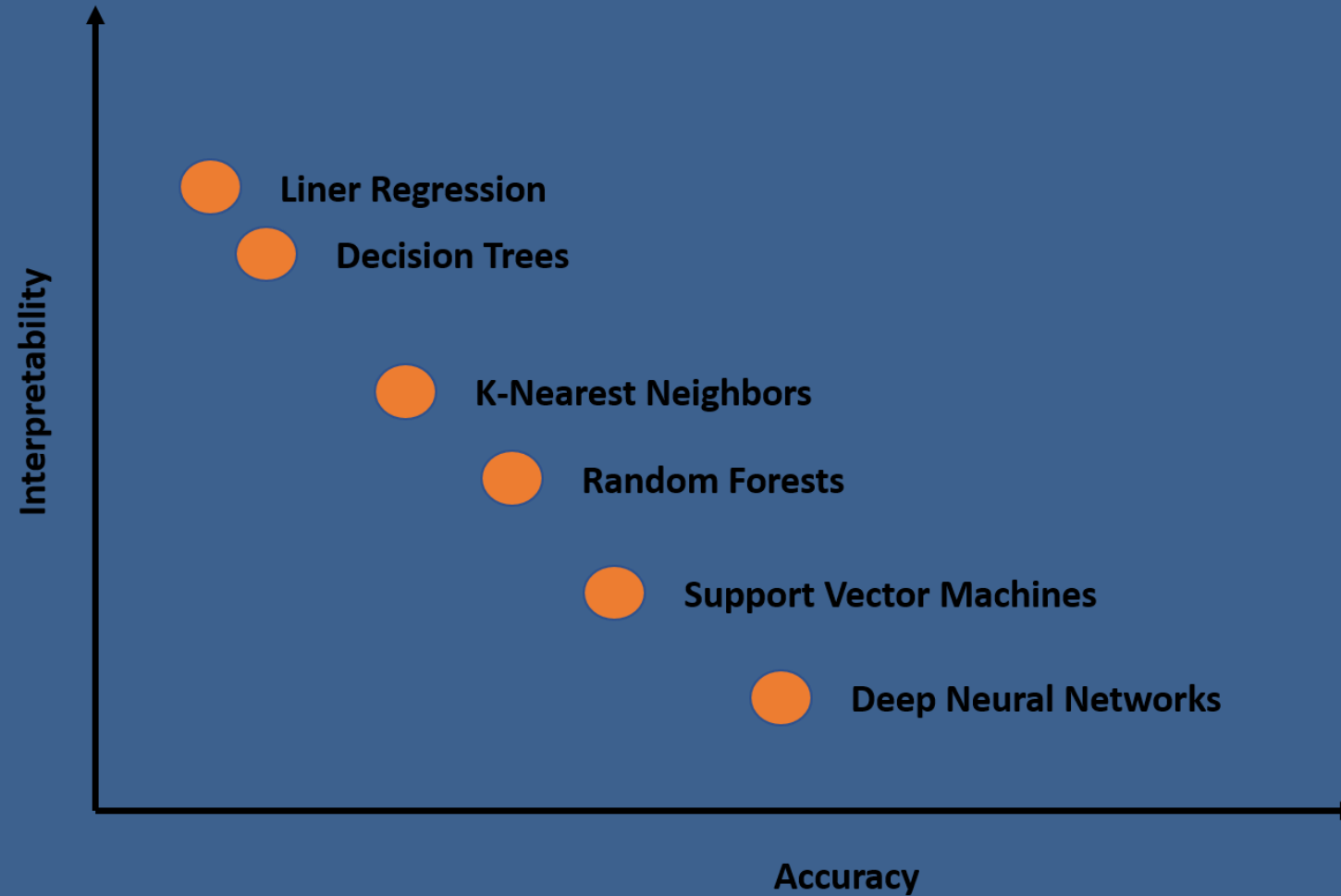








# But...



(Application of artificial intelligence in gastroenterology, April 2019, World Journal of Gastroenterology 25(14):1666-1683)

# Explain-ability of AI Algorithms

Regression – Multinomial Equation can be messy

Random Forrest – Multiple Tree and their Data Set and Voting

SVM – Kernel and Data Partition effect on Feature

K Mean – Nature of Centroid don't describe cluster well.

NN – Hidden Nodes and their way of creating features

Explainable AI (XAI) – A Perspective, Saurabh Kaushik





So...Finally

**icertis**  
Applied Cloud

AI is just a set of algorithms,  
which (predominantly ) work given  
Inputs as well as Outputs



# For example: to find out clause “Term”...

Need to supply AI algorithm with 100s/1000s of “Term” clause samples

Gets trained on patterns, word frequencies, context words

Stores this information as ‘model’

Can be used to classify unseen clause, whether “Term” or not.

# Quiz

If you want to find something new say “LIBOR” clause, what would be needed?

# Let's Summarize together

AI-ML-DL approaches are not deterministic (but <>?)

For good results, need good annotated data and lots of it

Annotations need to be perfect ("Gold"), else Garbage-In-<??-<??>

Data should cover all possible variations

AI-ML-DL just fits the data, just that, but, it does it automatically!!

ML has better explain-ability than DL (why?)

# Btw, thoughts to ponder on...

'AI is biased' because of it's training/data

'Explainable AI' is to understand how/why it is biased

But then...

Human are biased too...

'Explainable Humans'... the next topic for talk? Anyone?

