Destination IP Virtual Summit

SLW INSTITUTE
Building Extraordinary IP Assets

Writing Patent Applications for Machine Learning and AI-Related Technologies

Before We Get Started...

Recording

A link to the recording and slides will be emailed to all registrants.

Questions

Type in the question box and we will answer in real time or during the Q&A.

Social

Follow us on LinkedIn or go to SLW Institute on slwip.com to see upcoming and on demand webinars.

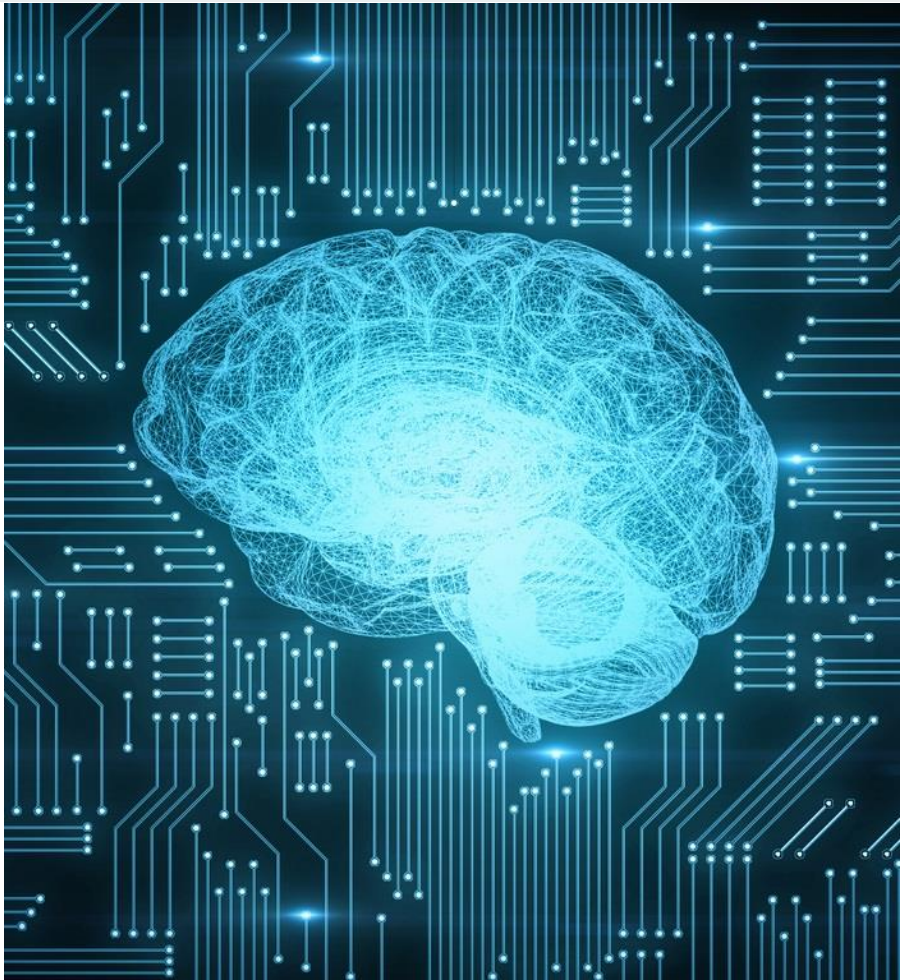
Panel

Sabine Volkmer Ward
Attorney

Jose Nunez
Attorney

Garth Vivier
Principal

- AI Patenting
- The Machine-Learning Framework
- Considerations for Drafting and Prosecution
- Machine-Learning Patent Examples
- Best Practices

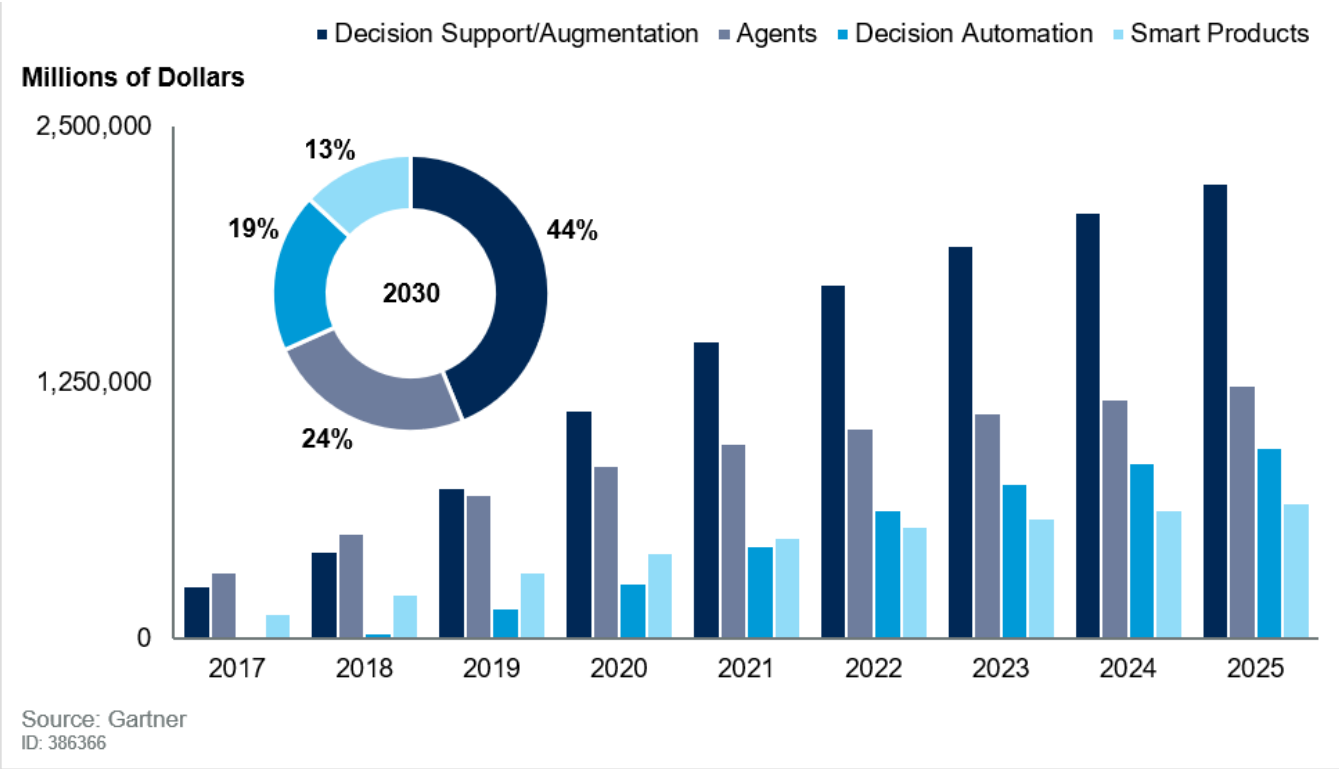


https://thispersondoesnotexist.com/

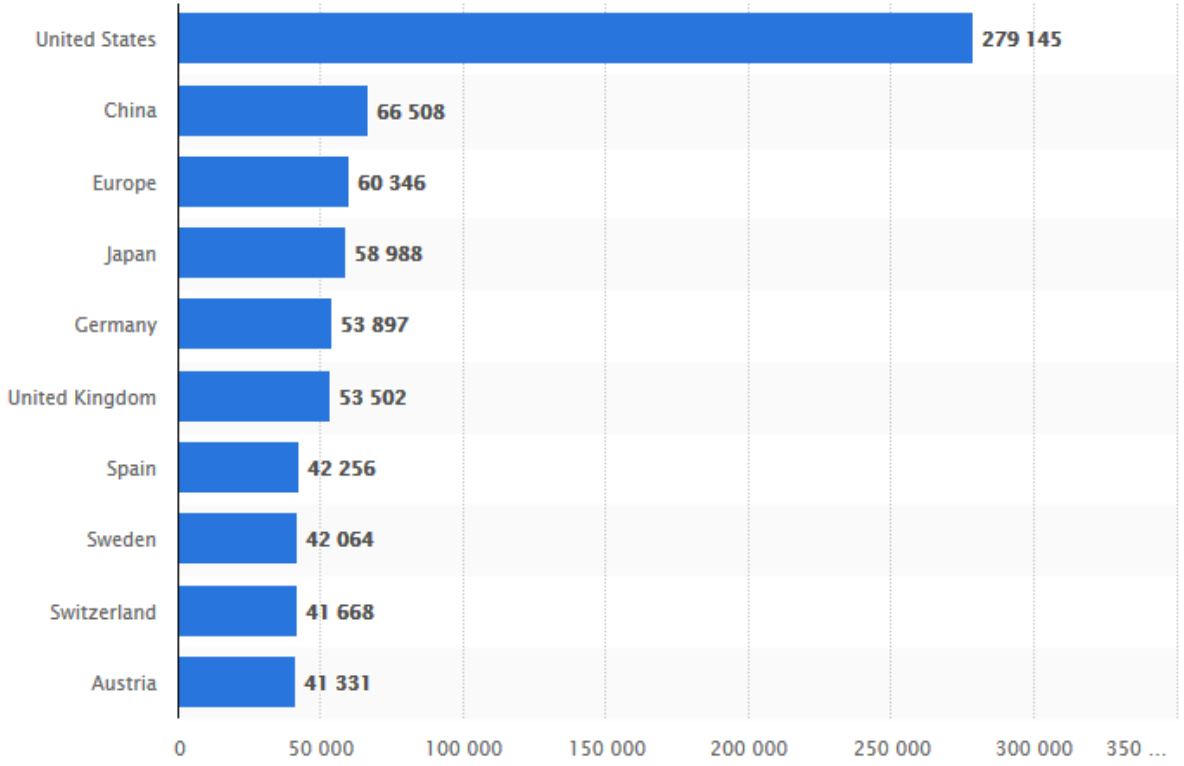
The site is the work of a team at NVIDIA, a major tech company specializing in artificial intelligence software. They trained a computer with 70,000 pictures of human faces pulled from Flickr. Then, the program was able to build entirely new faces from scratch by toggling different styles, colors and genders.

Business Value Forecast by AI Type

AI Augmentation Will Create $2.9 Trillion of Business Value in 2021

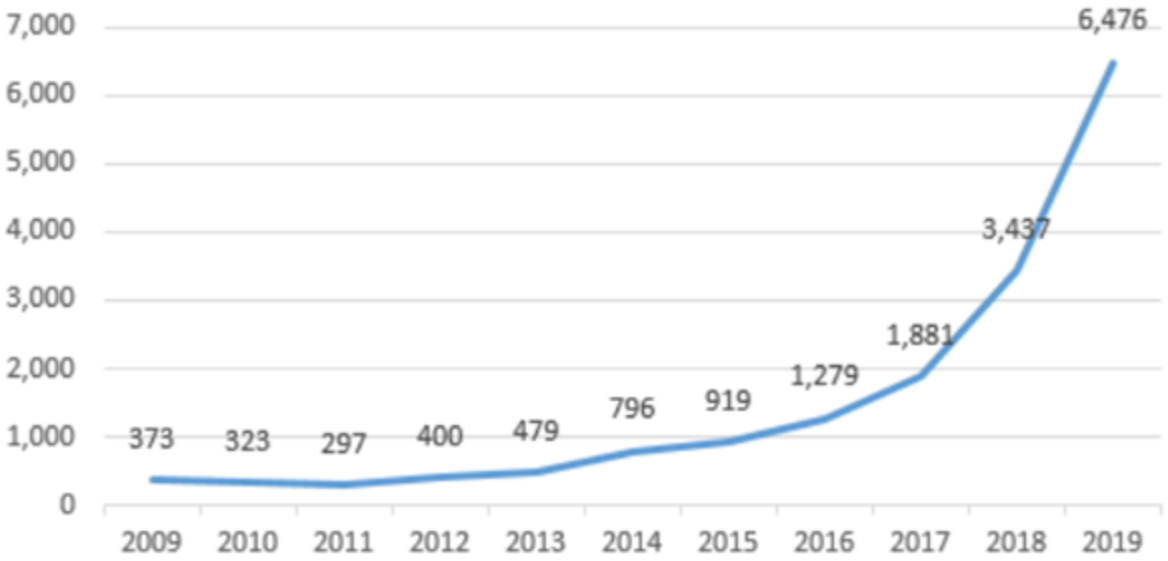


AI Patents By Geography



Published by Statista Research Department, March 2, 2020

AI Patent Publications in US



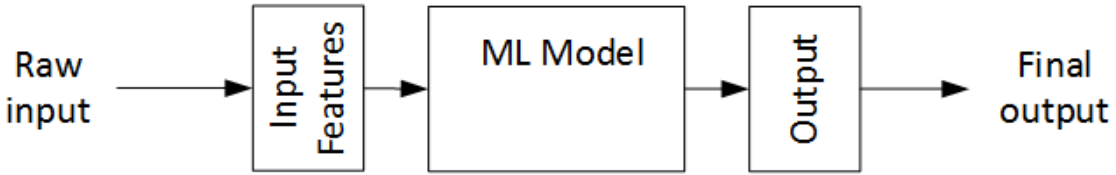
What is Artificial Intelligence

- The use of machines to mimic cognitive functions associated with humans, such as learning and problem solving
  - Techniques: Machine Learning, Logic Programming, Probabilistic Reasoning, Fuzzy Logic, …
- Law developing as AI developing
  - AI not an inventor
  - Moral issues
  - Ownership issues
- AI patents following same path as software patents
- Reasons to patent AI

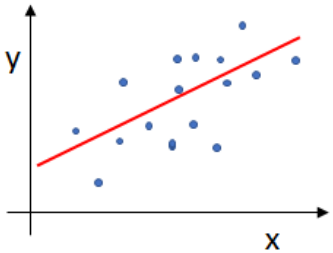
The ML Framework

- ML Models
- ML Algorithms to Build/Train Models
- Patentable Aspects of ML

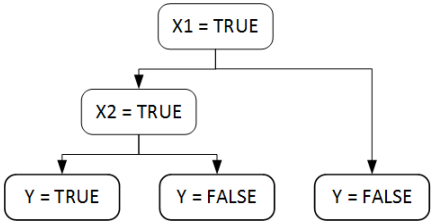
ML Models



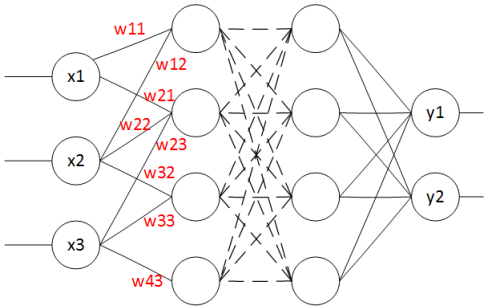
Regression



Decision tree



Neural network

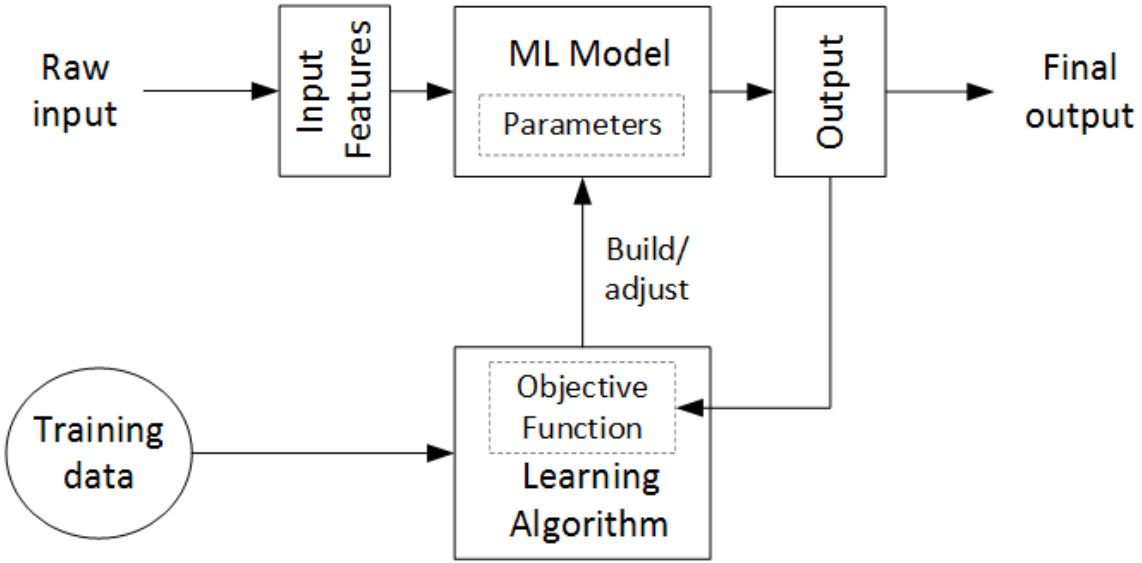


ML Algorithms

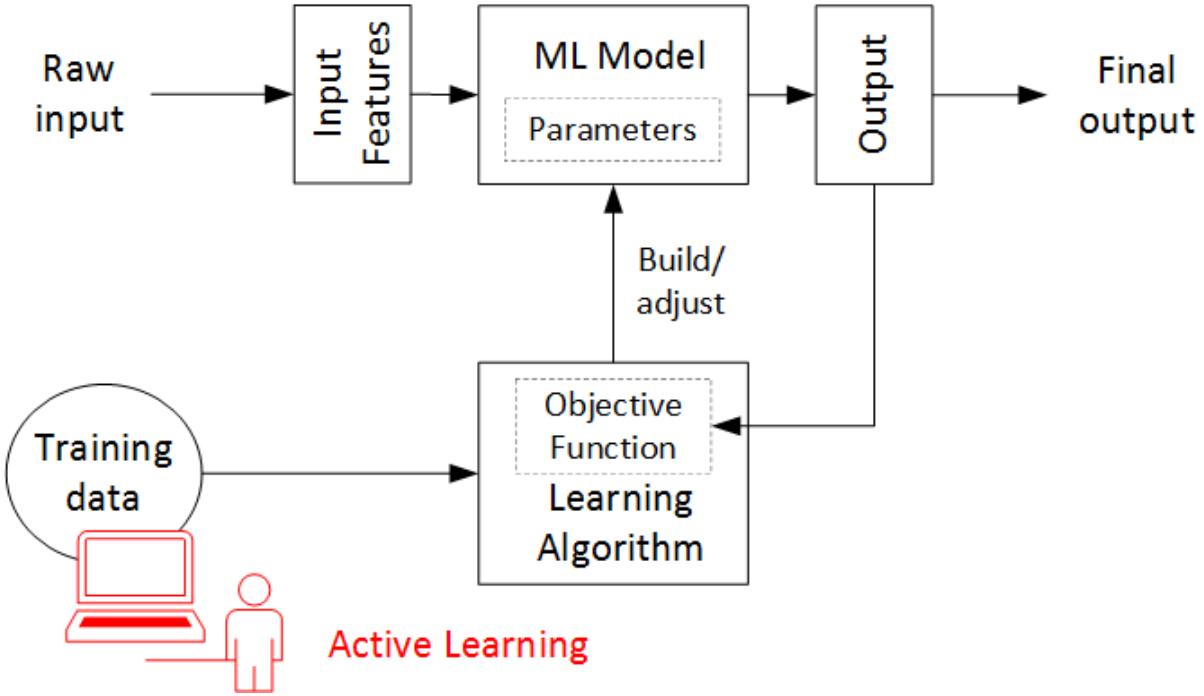
Regression
e.g., least squares

Decision tree
e.g., greedy recursive partitioning

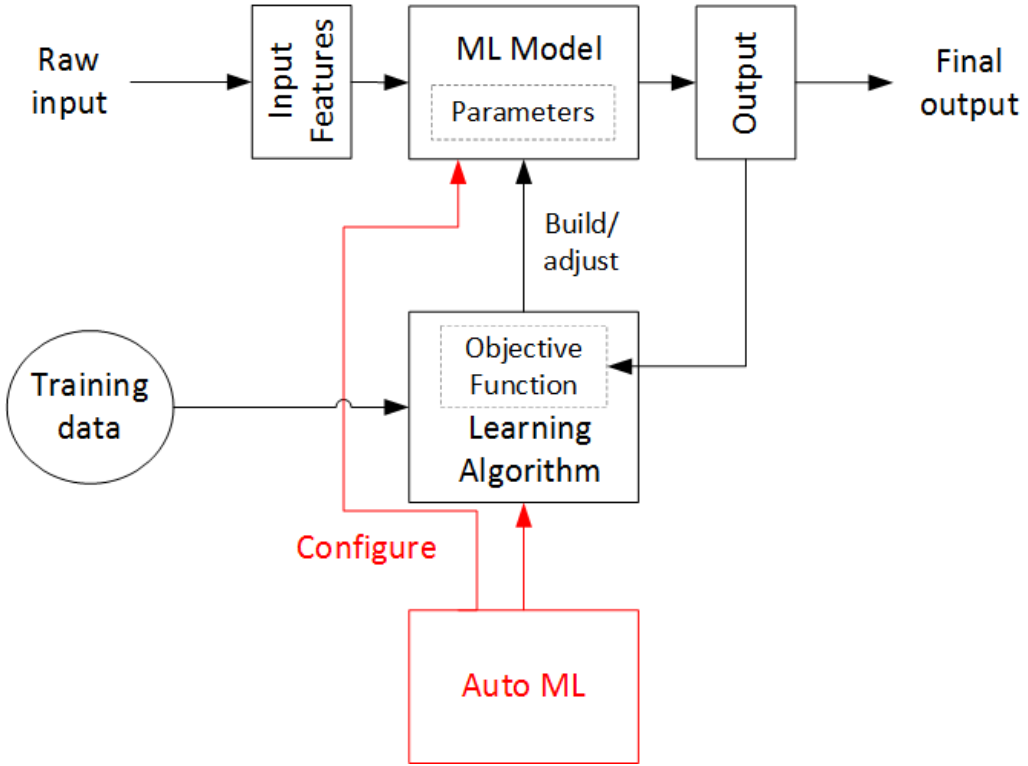
Neural network
e.g., backpropagation of errors



ML Tools


Raw input
Input Features
ML Model
Parameters
Output
Final output
Build/ adjust
Objective Function
Learning Algorithm
Training data
Active Learning

ML Tools


Raw input
Input Features
ML Model
Parameters
Output
Final output
Build/ adjust
Training data
Objective Function
Learning Algorithm
Configure
Auto ML

Drafting and Prosecuting Patent Applications for ML Inventions

- Understanding the Invention
- Describing the Invention
- Claiming the Invention
- Prosecution Obstacles

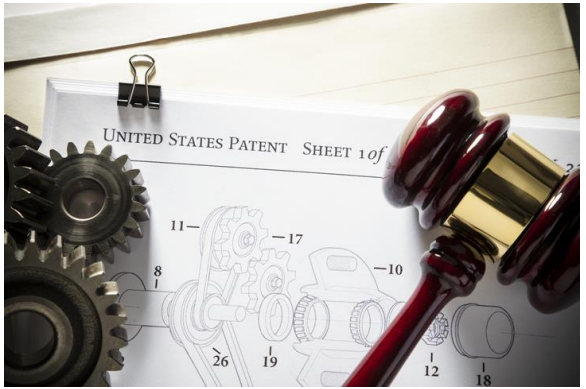
Understanding the Invention

- Learn basic ML and continue learning as you go
- Start with a good disclosure meeting and follow up as needed
    - Inventors' academic articles can be useful
- Identify where the invention resides
- Sometimes you have to go down the rabbit hole

Describing the Invention

- Paint the big picture, but don't leave out the detail
  - Provide context and examples of applications
  - Discuss the technical problem/solution and/or technical improvements in the field
  - Use schematics to illustrate basic concepts
  - Don't be afraid of math
- Focus on the point of novelty, but provide enough context for full enablement
- Reference known ML concepts and/or use a good boilerplate (often useful during prosecution)

Claiming the Invention



- Don't leave the invention out of the claim
- Claim the "how to," not just the "what"
- Tie in practical application
    - Balance the trade-off between breadth and patentability (sweet spot may differ between jurisdictions)
- Consider detectability and infringement scenarios
    - Training vs. inference

Prosecution Obstacles

- Anything can be a model
  - Learning from data, not heuristics
- Done with pen and paper
  - No “machine learning” without “machine”
- Mention of ML in reference implies all applications of ML are covered
  - Show distinguishing technical details
  - Show technical details in the claims
  - Focus on novelty
- From personal experience, it appears there are fewer 101 rejections lately

Sample ML Patents

- Training Process
- Use of Model/Application
- Preparing Data for ML
- Neural Network Model and Training Input

Adversarial Teacher-Student Learning for Unsupervised Domain Adaptation

Claiming the Training Process

US 20190287515A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2019/0287515 A1**

Li et al. (43) **Pub. Date: Sep. 19, 2019**

(54) **ADVERSARIAL TEACHER-STUDENT LEARNING FOR UNSUPERVISED DOMAIN ADAPTATION**

(71) Applicant: **Microsoft Technology Licensing, LLC**, Redmond, WA (US)

(72) Inventors: **Jinyu Li**, Redmond, WA (US); **Zhong Meng**, Redmond, WA (US); **Yifan Gong**, Sammamish, WA (US)

(21) Appl. No.: **15/923,795**

(22) Filed: **Mar. 16, 2018**

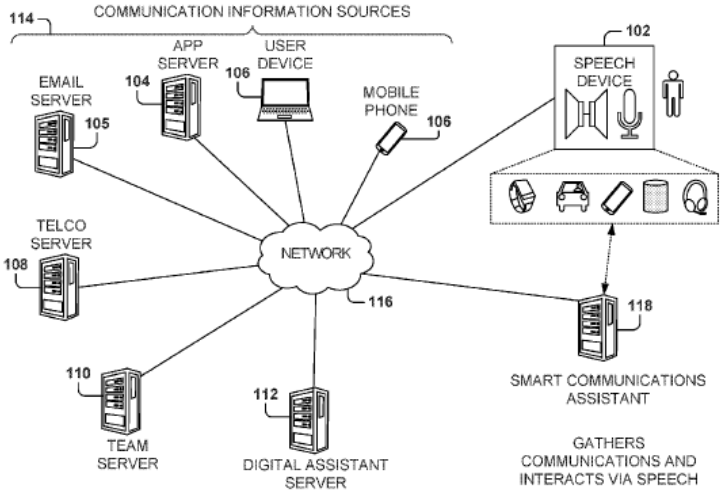
Publication Classification

(51) **Int. Cl.**
G10L 15/06 (2006.01)
G06N 99/00 (2006.01)

(52) **U.S. Cl.**
CPC ***G10L 15/063*** (2013.01); ***G06N 99/005*** (2013.01)

(57) **ABSTRACT**

Methods, systems, and computer programs are presented for training, with adversarial constraints, a student model for speech recognition based on a teacher model. One method includes operations for training a teacher model based on teacher speech data, initializing a student model with parameters obtained from the teacher model, and training the student model with adversarial teacher-student learning based on the teacher speech data and student speech data. Training the student model with adversarial teacher-student learning further includes minimizing a teacher-student loss that measures a divergence of outputs between the teacher model and the student model; minimizing a classifier condition loss with respect to parameters of a condition classifier; and maximizing the classifier condition loss with respect to parameters of a feature extractor. The classifier condition loss measures errors caused by acoustic condition classification. Further, speech is recognized with the trained student model.



A Student Model Leverages a Teacher Model

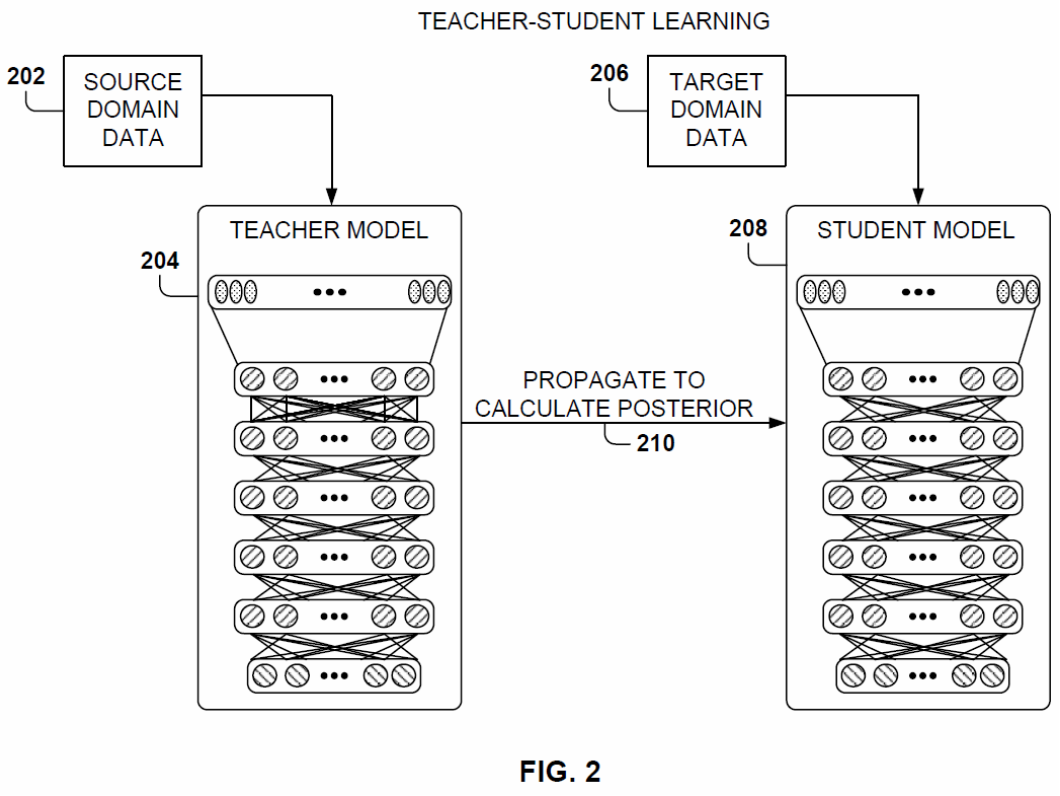
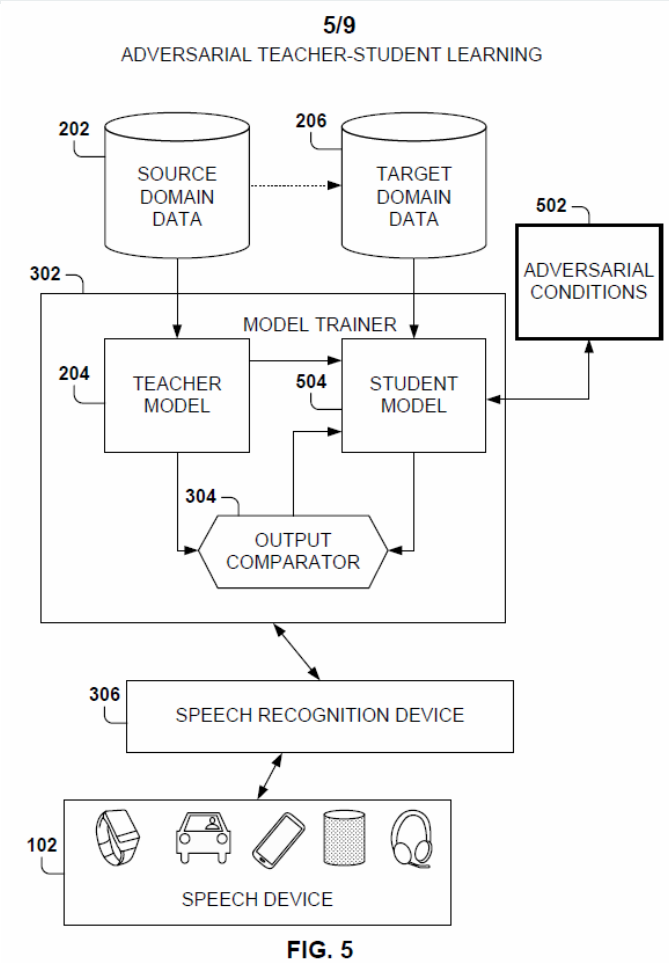


FIG. 2

Adversarial Conditions



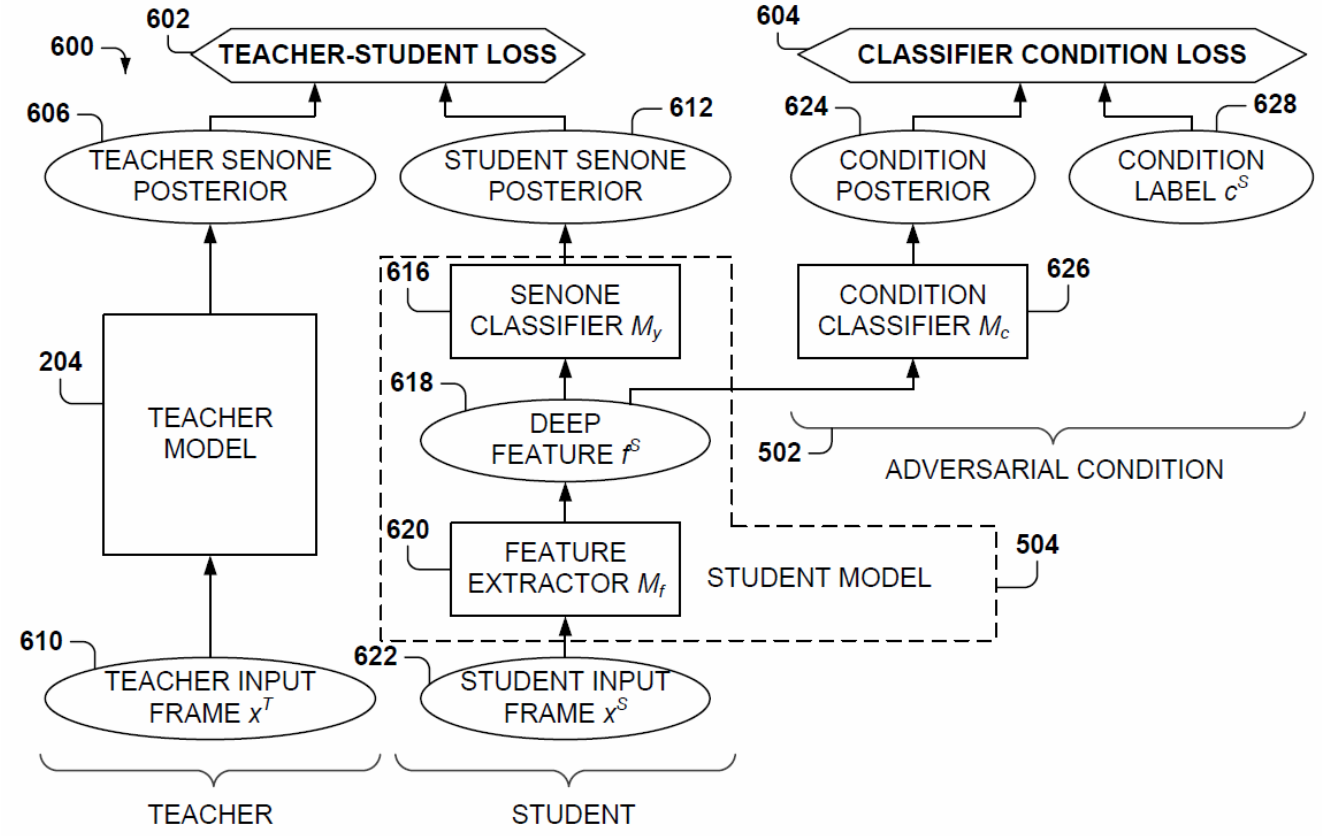

600
602
TEACHER-STUDENT LOSS
604
CLASSIFIER CONDITION LOSS
606
TEACHER SENONE POSTERIOR
612
STUDENT SENONE POSTERIOR
624
CONDITION POSTERIOR
628
CONDITION LABEL c^S
616
SENONE CLASSIFIER M_y
626
CONDITION CLASSIFIER M_c
204
TEACHER MODEL
618
DEEP FEATURE f^S
502
ADVERSARIAL CONDITION
620
FEATURE EXTRACTOR M_f
STUDENT MODEL
504
610
TEACHER INPUT FRAME x^T
622
STUDENT INPUT FRAME x^S
TEACHER
STUDENT

FIG. 6

1. A method comprising:

training, by one or more processors, a teacher model based on teacher speech data;

initializing, by the one or more processors, a student model with parameters obtained from the trained teacher model;

training, by the one or more processors, the student model with adversarial teacher-student learning based on the teacher speech data and student speech data, training the student model with adversarial teacher-student learning further comprising:

minimizing a teacher-student loss that measures a divergence of outputs between the teacher model and the student model;

minimizing a classifier condition loss with respect to parameters of a condition classifier, the classifier condition loss measuring errors caused by acoustic condition classification; and

maximizing the classifier condition loss with respect to parameters of a feature extractor; and

recognizing speech with the trained student model.

Allowed without prosecution

Smart communications assistant with audio interface

Claiming the Application

(12) **United States Patent**
Ghotbi et al.

(10) **Patent No.: US 10,516,637 B2**
(45) **Date of Patent: Dec. 24, 2019**

(54) **SMART COMMUNICATIONS ASSISTANT WITH AUDIO INTERFACE**

(71) Applicant: **Microsoft Technology Licensing, LLC**, Redmond, WA (US)

(72) Inventors: **Nikrouz Ghotbi**, Redmond, WA (US); **August Nichaus**, Redmond, WA (US); **Sachin Venugopalan**, Redmond, WA (US); **Aleksandar Antonijevic**, Bellevue, WA (US); **Tvrtko Tadic**, Seattle, WA (US); **Vashutosh Agrawal**, Bellevue, WA (US); **Lisa Stifelman**, Palo Alto, CA (US)

(73) Assignee: **Microsoft Technology Licensing, LLC**, Redmond, WA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 80 days.

(21) Appl. No.: **15/786,184**

(22) Filed: **Oct. 17, 2017**

(65) **Prior Publication Data**

US 2019/0116144 A1 Apr. 18, 2019

(51) **Int. Cl.**
G06F 15/16 (2006.01)
H04L 12/58 (2006.01)
(Continued)

(52) **U.S. Cl.**
CPC ***H04L 51/063*** (2013.01); ***G06F 3/167*** (2013.01); ***G06F 16/345*** (2019.01);
(Continued)

(58) **Field of Classification Search**
CPC H04L 51/063; H04L 51/046; H04L 51/26; G06F 3/167; G06F 16/345; G06F 17/2785; G06N 5/02; G10L 15/22
(Continued)

(56) **References Cited**

U.S. PATENT DOCUMENTS

7,415,100 B2 * 8/2008 Cooper H04M 3/527 379/88.01
10,186,267 B1 * 1/2019 Devaraj H04L 67/306
(Continued)

FOREIGN PATENT DOCUMENTS

WO 2017200595 A1 11/2017

OTHER PUBLICATIONS

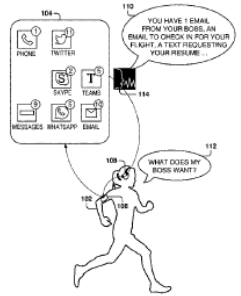
"International Search Report and Written Opinion Issued in PCT Application No. PCT/US2018/055119", dated Dec. 7, 2018, 11 Pages.

Primary Examiner — Kristie D Shingles
(74) *Attorney, Agent, or Firm* — Jeffrey Ranck; Ranck IP Law

(57) **ABSTRACT**

Methods, systems, and computer programs are presented for a smart communications assistant with an audio interface. One method includes an operation for getting messages addressed to a user. The messages are from one or more message sources and each message comprising message data that includes text. The method further includes operations for analyzing the message data to determine a meaning of each message, for generating a score for each message based on the respective message data and the meaning of the message, and for generating a textual summary for the messages based on the message scores and the meaning of the messages. A speech summary is created based on the textual summary and the speech summary is then sent to a speaker associated with the user. The audio interface further allows the user to verbally request actions for the messages.

19 Claims, 9 Drawing Sheets



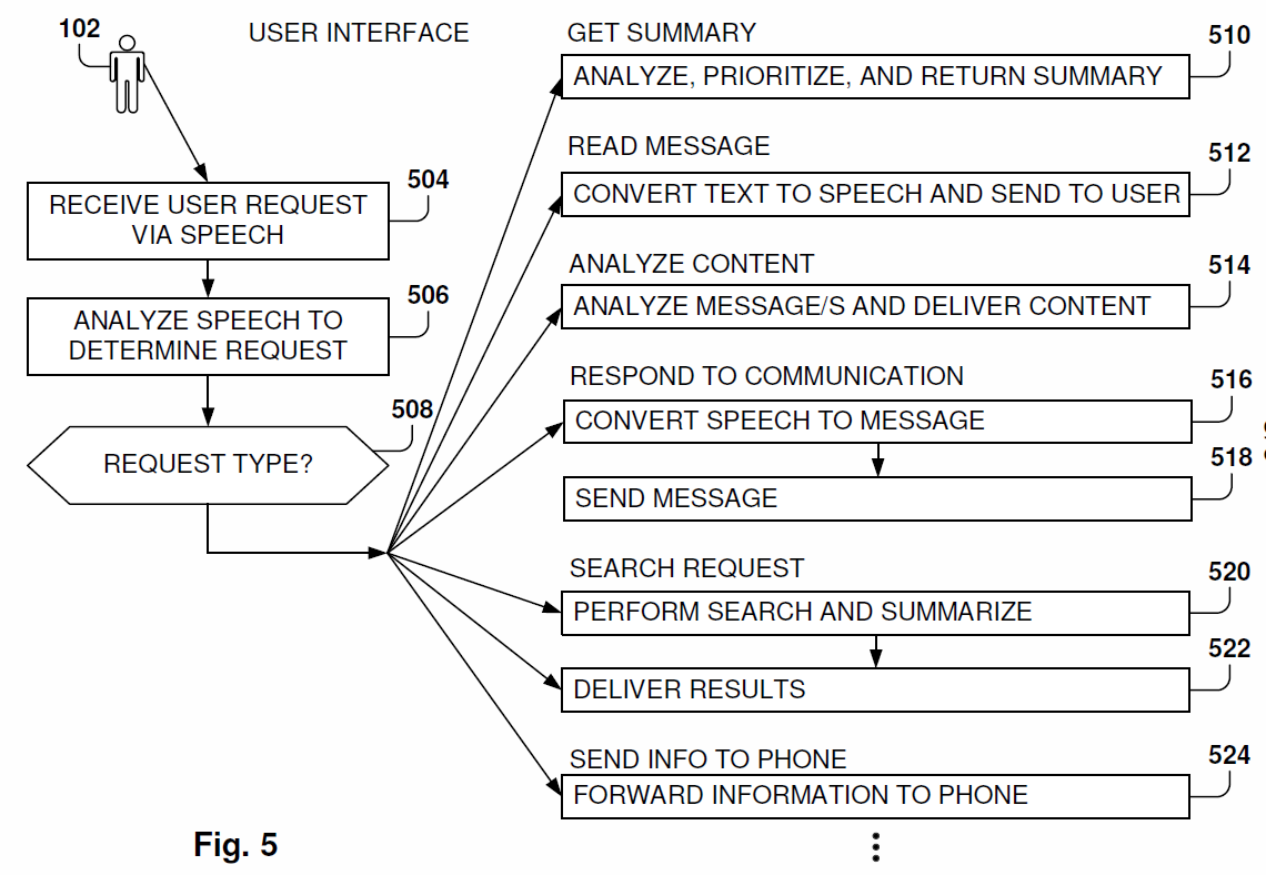

102
USER INTERFACE
RECEIVE USER REQUEST VIA SPEECH
504
ANALYZE SPEECH TO DETERMINE REQUEST
506
REQUEST TYPE?
508
GET SUMMARY
ANALYZE, PRIORITIZE, AND RETURN SUMMARY
510
READ MESSAGE
CONVERT TEXT TO SPEECH AND SEND TO USER
512
ANALYZE CONTENT
ANALYZE MESSAGE/S AND DELIVER CONTENT
514
RESPOND TO COMMUNICATION
CONVERT SPEECH TO MESSAGE
516
SEND MESSAGE
518
SEARCH REQUEST
PERFORM SEARCH AND SUMMARIZE
520
DELIVER RESULTS
522
SEND INFO TO PHONE
FORWARD INFORMATION TO PHONE
524

Fig. 5

Combining Multiple Models

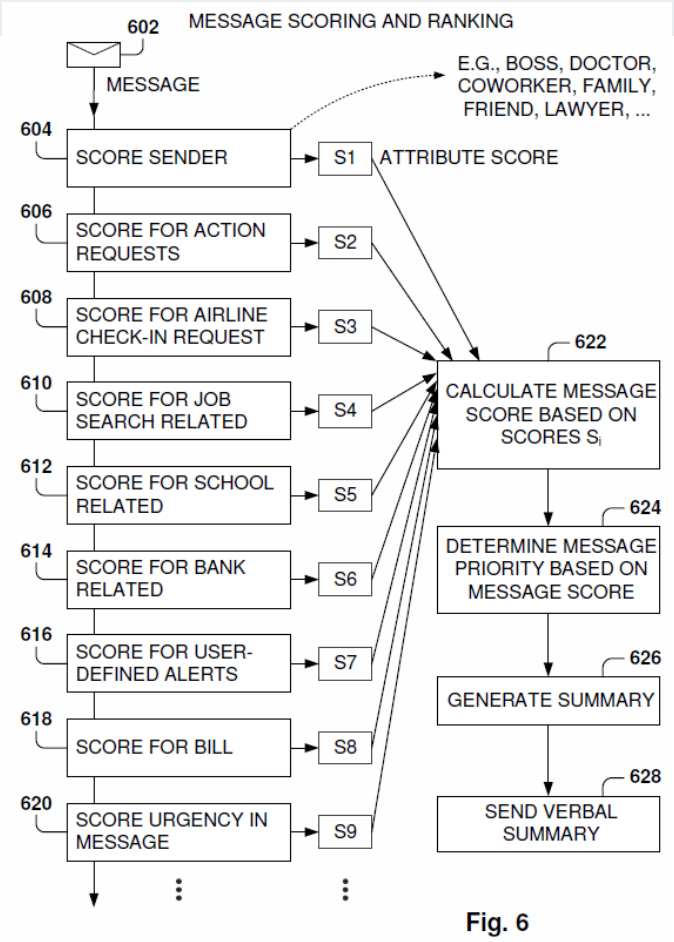


Fig. 6

Describing Features for ML

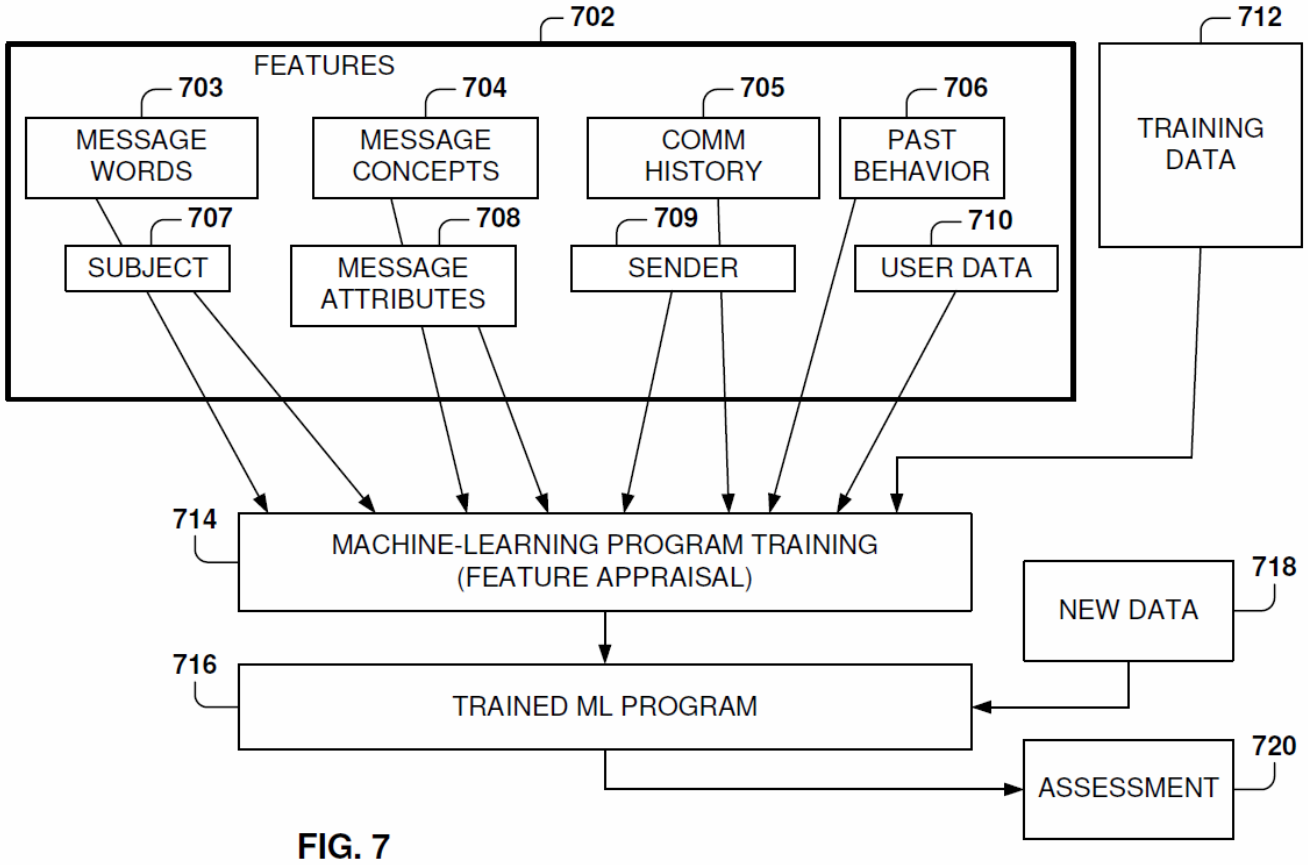


FIG. 7

Claims

1. A method comprising:

getting, by one or more processors, one or more messages from one or more message sources, the one or more messages being addressed to a user, each message comprising message data that includes message text;

analyzing, by the one or more processors, the message data to determine a meaning of the message text in each message;

generating, by the one or more processors, a score for each message based on the respective message data and the meaning of the message;

selecting messages based on the score of each message;

generating, by the one or more processors, a textual summary for the selected messages based on the meaning of the message text;

creating, by the one or more processors, a speech summary for each selected message, based on the textual summary and the meaning of the message text; and

sending, by the one or more processors, the speech summary for the selected messages to a speaker associated with the user.

Image recognition with promotion of underrepresented classes

Preparing the Data For ML

(12) **United States Patent**
Guo et al.

(10) **Patent No.: US 10,546,232 B2**
(45) **Date of Patent: Jan. 28, 2020**

(54) **IMAGE RECOGNITION WITH PROMOTION OF UNDERREPRESENTED CLASSES**

(71) Applicant: **Microsoft Technology Licensing, LLC**, Redmond, WA (US)

(72) Inventors: **Yandong Guo**, Bellevue, WA (US); **Lei Zhang**, Bellevue, WA (US)

(73) Assignee: **Microsoft Technology Licensing, LLC**, Redmond, WA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 302 days.

(21) Appl. No.: **15/722,821**

(22) Filed: **Oct. 2, 2017**

(65) **Prior Publication Data**

US 2019/0012526 A1 Jan. 10, 2019

Related U.S. Application Data

(60) Provisional application No. 62/528,508, filed on Jul. 4, 2017.

(51) **Int. Cl.**
G06N 3/04 (2006.01)

(52) **U.S. Cl.**
CPC ***G06N 3/0454*** (2013.01)

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**

FOREIGN PATENT DOCUMENTS

CN 103778412 A 5/2014

OTHER PUBLICATIONS

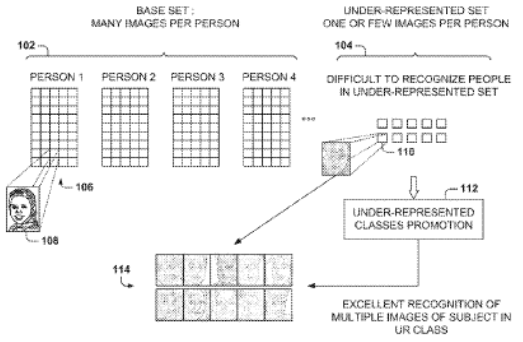
Pemula, Latha. "Low-shot Visual Recognition." PhD diss., Virginia Tech, 2016. (Year: 2016).*
Hanselmann, Harald, Shen Yan, and Hermann Ney. "Deep fisher faces." In BMVC. 2017. (Year: 2017).*
Wen, Yandong, Kaipeng Zhang, Zhifeng Li, and Yu Qiao. "A discriminative feature learning approach for deep face recognition." In European conference on computer vision, pp. 499-515. Springer, Cham, 2016. (Year: 2016).*
Deng, et al., "Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach", In Journal of Pattern Recognition, vol. 43, Issue 5, May 31, 2010, pp. 1748-1762.
(Continued)

Primary Examiner — David Perlman
(74) *Attorney, Agent, or Firm* — Schwegman Lundberg & Woessner, P.A.

(57) **ABSTRACT**

Methods, systems, and computer programs are presented for a classifier that recognizes images when at least one class includes just a few training samples. One method includes identifying a training set containing samples, each sample associated with a class from many classes. Further, the method divides the training set into a base and a novel set based on the number of samples in each class, trains a first classifier with the base set, and trains a second classifier using the training set. The second classifier is trained with promotion of the novel set and based on minimizing a loss function that comprises a first term and a second term, the first term associated with a first summation for the samples of the training set, the second term associated with a second summation for the samples of the novel set. Further, the method classifies an item with the trained second classifier.

20 Claims, 8 Drawing Sheets



Problem with Insufficient Training Data

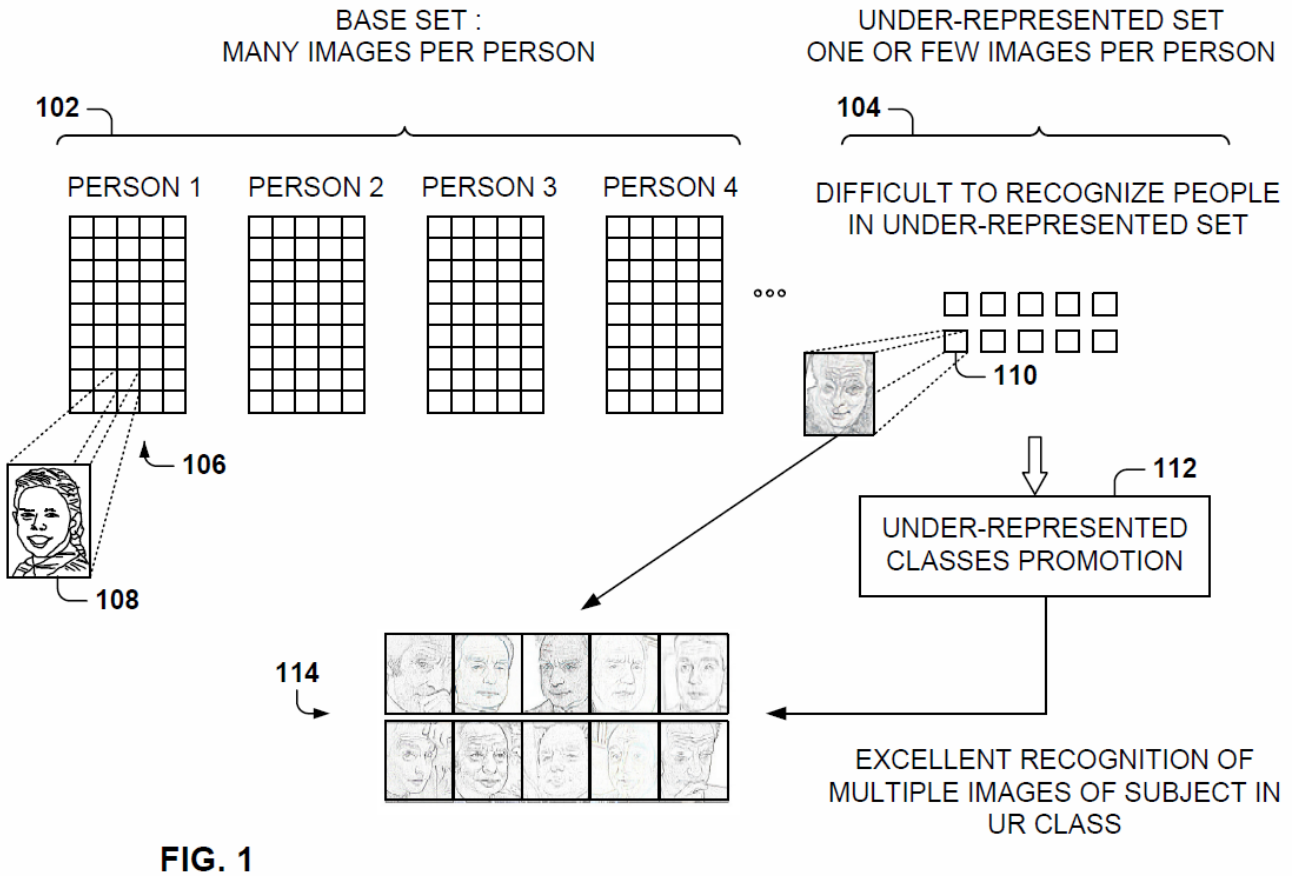


FIG. 1

Promote the Underrepresented Class

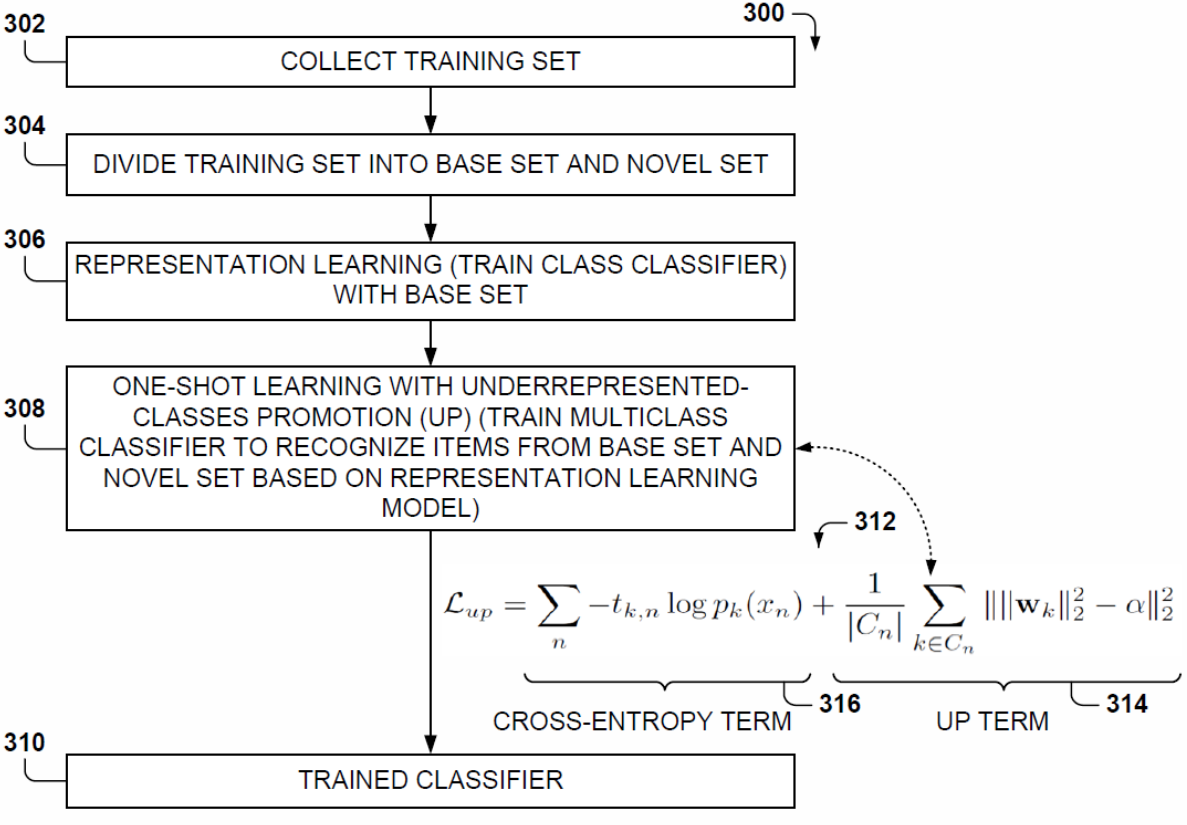


FIG. 3

Promotion Creates Space for UR Class

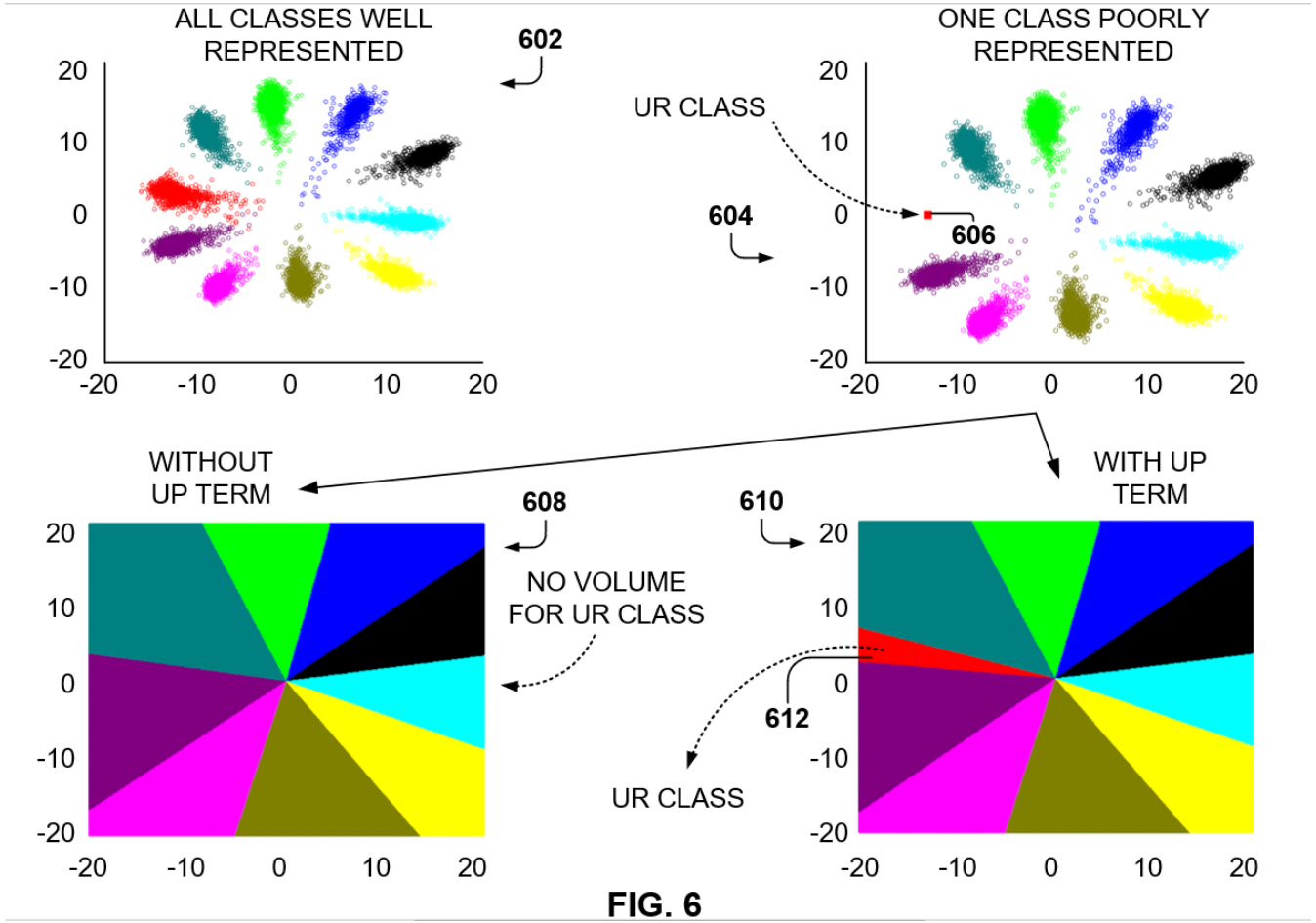


FIG. 6

1. A method comprising:

identifying, by one or more processors, a training set comprising a plurality of samples, each sample of the training set being associated with a class from a plurality of classes;

dividing, by the one or more processors, the training set into a base set and a novel set based on a number of samples in each class;

training, by the one or more processors, a first classifier with the base set and without the novel set;

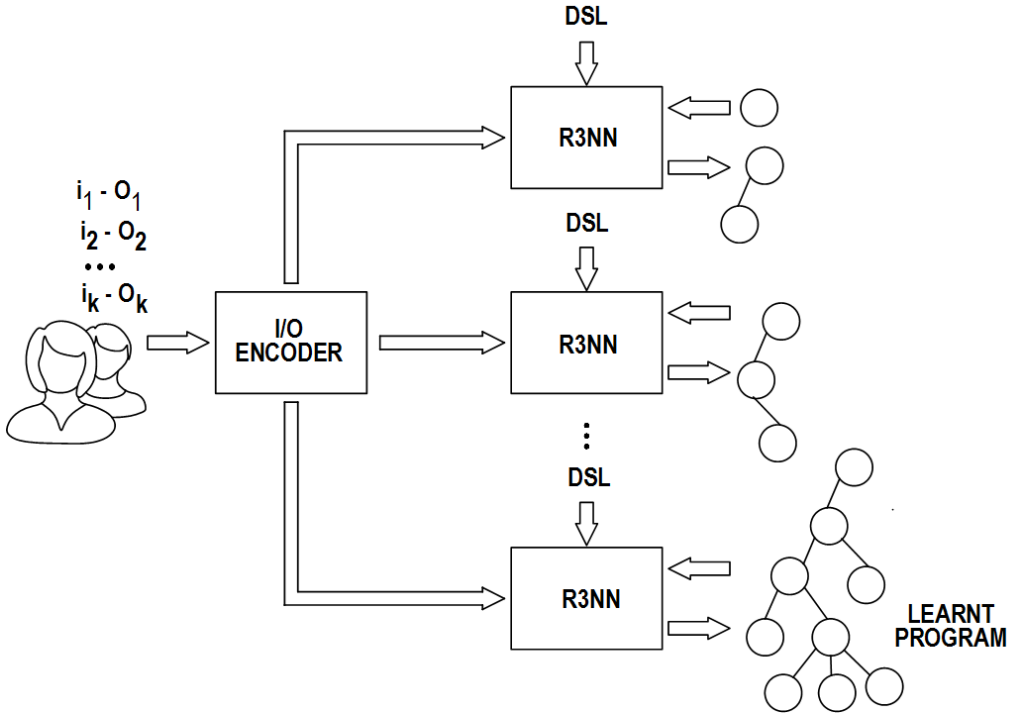
training, by the one or more processors, a second classifier, using the training set, with promotion of the novel set, the training of the second classifier being based on minimizing a loss function that comprises a first term and a second term, the first term being associated with a first summation for the samples of the training set, the second term being associated with a second summation for the samples of the novel set; and

classifying, by the one or more processors, an item with the trained second classifier.

Allowed without prosecution

ML for Program Synthesis – Training a Neural Network to Write Programs

INPUT	OUTPUT
William Henry Charles	Charles, W.
Michael Johnson	Johnson, M.
Barack Rogers	Rogers, B.
Martha D. Saunders	Saunders, M.
Peter T Gates	Gates, P.



1. A method comprising:

for a given domain-specific language that defines a plurality of symbols and a plurality of production rules, providing an input-output encoder and a program-generation model comprising a neural network, the input-output encoder and the neural network ***having been trained*** on a plurality of programs within the domain-specific language and a plurality of respective training sets of input-output examples associated with the programs, wherein, for each of the plurality of programs and its associated training set, each input-output example of the training set comprises a pair of an input to the program and a corresponding output produced by the program from the input;

providing a test set of input-output examples for a target program;

using one or more hardware processors to perform operations for generating the target program based on the test set of input-output examples, the operations comprising:

encoding the test set of input-output examples using the input-output encoder;

conditioning the program-generation model on the encoded set of input-output examples; and

using the neural network to generate a program tree representing the target program by iteratively expanding a partial program tree, beginning with a root node and ending when all leaf nodes are terminal, based on a computed probability distribution for a set of valid expansions, wherein leaves in the program tree and the partial program tree represent symbols in the domain-specific language and wherein non-leaf interior nodes in the program tree and the partial program tree represent production rules in the domain-specific language.

3. The method of claim 2,

wherein the recursive-reverse-recursive neural network specifies distributed representations of the plurality of symbols and the plurality of production rules and, for each of the plurality of production rules, first and second deep neural networks, and

wherein iteratively expanding the partial program tree comprises, in each iteration:

computing global leaf representations for at least non-terminal ones of the leaf nodes of the partial program tree by retrieving the distributed representations of the symbols represented by the leaf nodes, performing a recursive bottom-to-top pass through the partial program tree from the leaf nodes to the root node using the first deep neural networks, and thereafter performing a reverse-recursive top-to-bottom pass through the partial program tree from the root node to the leaf nodes using the second deep neural networks;

computing the probability distribution for the set of valid expansions from the global leaf representations and the distributed representations of the production rules;

selecting a non-terminal leaf node and a production rule based on the computed probability distribution; and

expanding the partial program tree by applying the selected production rule to the selected non-terminal leaf node.

Best Practices

- Identify the point of novelty – go down the rabbit whole if you need
- Capture the basic idea in words and drawings, but include formulas as back-up
- Vary claim scope surrounding point of novelty
- Describe both practical applications and technical benefits
- Consult counsel in other jurisdictions on patentability and adapt application

Thank you for your interest.

Questions?

Garth Vivier – gvivier@slwip.com
Sabine Volkmer Ward – sward@slwip.com
Jose Nunez – jnunez@slwip.com

SLW INSTITUTE
Building Extraordinary IP Assets

Subject Matter Eligibility Considerations

- Law pertaining to software and mathematical algorithms still evolving
    - 2019 Revised Subject Matter Eligibility Guidance from USPTO has been helpful

      "A claim that integrates a judicial exception into a ***practical application*** will apply, rely on, or use the judicial exception in a manner that imposes a ***meaningful limit*** on the judicial exception, such that the claim is more than a drafting effort designed to monopolize the judicial exception."
- Include the practical application in the claims
- Claim the "how to," not just the "what" (negates preemption)
- Argue improvements within a technical field

Subject Matter Eligibility in Europe

- Guidelines for Examination in EPO treat AI/ML like other computational models and algorithms
- Step 1: Invention has a ***technical character*** as a whole if it is directed to a method involving the ***use of technical means*** (e.g., computer) or to the device itself
- Step 2: Assessment of inventive step involves asking whether the ***mathematical method contributes to the technical character***
    - Adapted for a specific ***technical implementation*** (internal functioning of the computer)
    - Serves a specific ***technical purpose***

Subject-Matter Eligibility in Europe

Examples of technical applications

- Controlling a specific technical system or process (e.g., X-ray apparatus)
- Audio, image, or video analysis/enhancement
- Speech recognition
- Audio, image, video, sensor data compression
- Encryption or signing of electronic messages
- Automated medical diagnosis by processing physiological measurements (caveat: not physician's actions)

Problematic areas

- Programming
- Processing text (considered "linguistic purpose")

These materials are for general informational purposes only. They are not intended to be legal advice, and should not be taken as legal advice. They do not establish an attorney-client relationship.

SLW INSTITUTE
Building Extraordinary IP Assets