Soup to Nuts of Protecting ML Innovations

Sumedha Ahuja
Partner
Perkins Coie LLP
Today’s Presenters

Sumedha Ahuja
Partner, Perkins Coie LLP

sahuja@perkinscoie.com
I. What is AI/ML
II. Conducting effective disclosure meetings for ML innovations
III. Challenges to patenting AI
   1. Subject Matter Eligibility (Section 101)
   2. Written Description and Enablement (Section 112)
   3. Prior-art based rejections (Section 103)
AI is Everywhere
Many different definitions and subject to change in the future

- “Software and/or hardware that can learn to solve complex problems, ... undertake tasks that require human-like ..., cognition, planning, learning, communication, or physical action” (NIST)

ML = Automatically deriving useful signals from data

Source: Office of the Chief Economist, “Inventing AI”, Number 5, October 2020
Comparing ML to Traditional Software

Traditional Programming

Machine Learning
Pipeline View of an ML System

DATA SOURCING
- Data Filtering, Scrubbing, Normalization

MODEL ORGANIZATION AND TRAINING

MODEL VALIDATION

MODEL APPLICATION

INPUT

OUTPUT

Runtime/Execution Pipeline

Model Design

Model Implementation

Parameters

Training Pipeline

INPUT

OUTPUT
How to Conduct an Effective Disclosure Meeting

1. Location of invention
2. Problem being solved
3. Data collection and pre-processing
4. Model architecture and training
5. Post-processing steps
6. Output utilization
With Whom to Conduct the Disclosure Meeting

• PARTICIPANTS
  - High-level technical person – CTO, CDO, Chief Scientist
  - Mid-level technical person in key areas
  - Sales person most responsible for product/product category
Draft Competitively

Assess competitive value when prioritizing ML pipeline inventions:

- **Prioritize “leverage” technology** that is (or will be) necessary for a competitor to compete with you
- Carefully consider substitutions/design-around potential within the ML pipeline
- **Evaluate and claim unavoidable requirements of competitor** ML pipeline solutions (even if different from your own solution)
- Include substantial discussion of the practical application(s) (i.e., what you enable/accomplish by using the ML pipeline)
Patenting AI – Legal Challenges

For patent protection, an invention

- **statutory** (35 U.S.C. § 101)
- new
- useful
- **non-obvious** (35 U.S.C. §§ 102, 103)

Global fights of naming AI system “DABUS” as inventor

AI ≠ “Black Box”

Other competitors using it?
Thank You!
Subject Matter Eligibility

• Common ML claim types:
  – Process (at any stage(s) of ML pipeline)
  – Structure (of neural network)
  – Data structure:
    • Trained model
    • Training observations
    • Scoring observations
Subject Matter Eligibility

• Where available, Machine Learning Model ("MLM") architecture claims may be less likely to be rejected than MLM training and application claims:
  – Network architecture: structure of single network
  – Macro-architecture: organization of multiple networks working together
  – Micro-architecture: new type of nodes or node-combinations
Subject Matter Eligibility – Certain Methods of Organizing Human Activity

<table>
<thead>
<tr>
<th>Ineligible</th>
<th>Generally Eligible</th>
<th>Practical Advice:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fundamental Economic Activities</td>
<td>Neural network architecture; applied classification systems</td>
<td>❑ <strong>Avoid end results</strong> and business advantages</td>
</tr>
<tr>
<td>Commercial or Legal Interactions</td>
<td>Data processing; applied learning</td>
<td>❑ <strong>Focus on technical aspects</strong> and benefits</td>
</tr>
<tr>
<td>Managing Behavior Relationships or Interactions</td>
<td>Autonomous vehicles; IoT; virtual assistants</td>
<td>❑ <strong>Target data normalization</strong>, mandatory/repetitive training requirements, and uses of the model</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ <strong>Discuss technical difficulties</strong> faced by existing technologies</td>
</tr>
<tr>
<td></td>
<td></td>
<td>❑ <strong>Describe practical or real-world applications</strong> of the claims with specificity</td>
</tr>
</tbody>
</table>

- Recommendation systems
- Productivity/workflow solutions
- Financial transactions
Subject Matter Eligibility – Mental Processes and Mathematical Concepts

Ineligible
- Observations, evaluations, judgements, opinions
- Bare formulas, equations, algorithms

Generally Eligible
- Steps incapable of being performed by “pen and paper”
- Applied formulas, equations, algorithms

Practical Advice:
- Avoid behavior and decision-making capable of being performed entirely in a human’s mind
- Explain necessity of a digital solution, and specify hardware components
- Generalize the ML model within the claim, to focus upon input/output novelty
- Patent Offices narrowly construe claimed algorithms (in general)
- Algorithmic claims may be necessary for certain inventions (e.g., codecs, standards, etc.)

- Signal processing
- Normalization (in the abstract)
- Weighing determination probabilities
- Weighing activation of nodes
- Solutions imitating/simulating human behavior
Written Description and Enablement

• Claimed aspects:
  – Very little guidance exists about the level of disclosure needed for machine learning claims recited training, storing, or applying a MLM
  – A conservative approach is to, where pursuing these claims, include such detail as:
    • model architecture—either a diagram, or an incorporation by reference of an article effectively describing a common architecture used
    • data dimensionality, and other details of training observation contents and organization
    • any nonstandard aspects of training scheme
Written Description and Enablement

• Aspects initially unclaimed:
  – It may be helpful to be able to add details during examination, particularly in response to eligibility rejections:
    • further detail about existing stage(s) of ML pipeline
    • detail about additional stage(s) of ML pipeline
  – Including well-crafted boilerplate content describing the entire ML pipeline at a reasonable level of detail can provide support for such amendments
Rejections under 35 U.S.C. 103 for claims reciting Machine-Learning
Barry J. Schindler  
Co-Chair, Global Patent Prosecution Group  
Greenberg Traurig, LLP  
500 Campus Drive, Suite 400  
Florham Park, NJ 07932  
Work: 973.360.7944  
Cell: 973.5194.944  
schindlerb@gtlaw.com

Lennie A. Bersh  
Shareholder  
Greenberg Traurig, LLP  
500 Campus Drive, Suite 400  
Florham Park, NJ 07932  
Work: 973.443.3536  
Cell: 212.767.9141  
bershl@gtlaw.com
Prima Facie Obviousness Based On ML Teachings: Rejections under 35 U.S.C. 103 for claims reciting Machine-Learning Models (MLMs)

- MPEP is silent
- Guiding Principles:
  - A ML Pipeline has particular objective(s) (e.g., solving a particular problem)
  - The Goal of MLM design, training and re-training is to satisfy such objective(s)
Guiding Principles (cont.):

Three key factors determine the accuracy of MLMs:

- **fit** (structure, inputs, outputs) of MLM: selecting a right MLM out of hundreds and hundreds of MLMs
- **completeness** of input data (e.g., feature vectors): the degree to which the number of data points required to reach a defined accuracy threshold has been provided
- **sufficiency** of training data
  - For example, if MLM is trained to identify cars from an image, and the input data consists of photos of airplanes, the model would not know what a car looks like. Such MLM will not provide good results.

- When a MLM consumes multiple inputs at once to predict multiple outputs at once – a relationship (e.g., index) between an order of inputs to an order of outputs is also critical to achieve a working MLM
- **Multiple-Model techniques require a defined relationship** between MLMs
Common *Prima Facie* Obviousness Rationales (MPEP 2143):

- **Combine** prior art MLMs according to known methods to yield *predictable* results; or
- **Substitute** of one prior art MLM for another prior art MLM to obtain *predictable* results; or
- **Modify** prior art MLM(s) “to arrive at the claimed invention” with “reasonable expectation of success”
**Prima Facie Obviousness Based On ML Teachings: Rejections under 35 U.S.C. 103 for claims reciting MLMs (cont.)**

- Meeting *prima facie based on a combination* of prior art MLMs according to known methods to yield predictable results:
  - Identify a known method on how to combine:
    - Output of MLM(1) as input for MLM(2); or
    - Apply a voting to outputs from different MLMs to obtain a common output; or
    - Apply a weighting function to numerical outputs from different MLMs to obtain a common score
  AND
  - Provide evidentiary support and/or technical reasoning as to why results would be predictable – e.g.:
    - How would the combination of prior art MLMs be trained to achieve an objective of either reference? or
    - In case when the combination is based on output of MLM(1) as input for MLM(2), how would output of MLM(1) meet the completeness for input data for MLM (2) and be related to output of MLM(2)?
Prima Facie Obviousness Based On ML Teachings: Rejections under 35 U.S.C. 103 for claims reciting MLMs (cont.)

- **Substitute** of MLM(1) for another MLM(2) to obtain predictable results:
  - Provide evidentiary support and/or technical reasoning for the substitution – e.g.:
    - **Why** would MLM(1) be fit to achieve the same **objective(s)** of MLM(2)? or
    - **Could** MLM(1) be **trained** with inputs of MLM(2)? – i.e., what is/are difference(s)/similarity(ies) between inputs of MLM(1) and MLM(2)?
Prima Facie Obviousness Based On ML Teachings: Rejections under 35 U.S.C. 103 for claims reciting MLMs (cont.)

- **Modify** a MLM “to arrive at the claimed invention” with “reasonable expectation of success”:
  - Provide evidentiary support and/or technical reasoning for the modification – e.g.:
    - Why would the modified MLM still fit so as to achieve the original objective(s)? or
    - How would the modified MLM be trained with original and/or modified input(s) to achieve the original objective(s)?
**Prima Facie Obviousness Based On ML Teachings: Rejections under 35 U.S.C. 103 for claims reciting MLMs (cont.)**

- **PTAB decisions:**
  - Examiner was **affirmed** when references in the substitution-based combination taught that their MLMs consumed **similar inputs** and were trained to achieve **similar objectives**
  - Examiner was **reversed** in the modification-based combination when the Board concluded that there was a lack of explanation as to why one would **modify** one MLM based on a **structure** of another MLM when those models provided **unrelated outputs** (i.e., models were designed for **unrelated objectives**).
**Prima Facie Obviousness Based On ML Teachings:**

Rejections under 35 U.S.C. 103 for claims reciting MLMs

**Examiners’ Obviousness “Toolbox”**

- **Combine** prior art MLMs according to known methods to yield **predictable** results; or
- **Substitute** of one prior art MLM for another prior art MLM to obtain **predictable** results; or
- **Modify** prior art MLM(s) “to arrive at the claimed invention” with “reasonable expectation of success”
Suggested Analytical Framework -- *Genus-Species* Analysis

- MPEP 2144.08: Obviousness of Species When Prior Art Teaches Genus
  
  “In the case of a prior art reference disclosing a genus, Office personnel should make findings as to:

  (A) the structure of the disclosed prior art genus and that of any expressly described species or subgenus within the genus;

  (B) any physical or chemical properties and utilities disclosed for the genus, as well as any suggested limitations on the usefulness of the genus, and any problems alleged to be addressed by the genus;

  (C) the predictability of the technology; and

  (D) the number of species encompassed by the genus taking into consideration all of the variables possible.”
Genus-Species Analysis (MPEP 2144.08) -- Not to re-invent a proverbial wheel

Potential factors to consider:

- the structure of disclosed prior art MLM genus and that of any expressly described MLM species or subgenus within the MLM genus;
- any similarity in design, training, and/or objective(s) addressed by the MLM genus/subgenus;
- the number of MLM species encompassed by the MLM genus taking into consideration all of the variables possible.