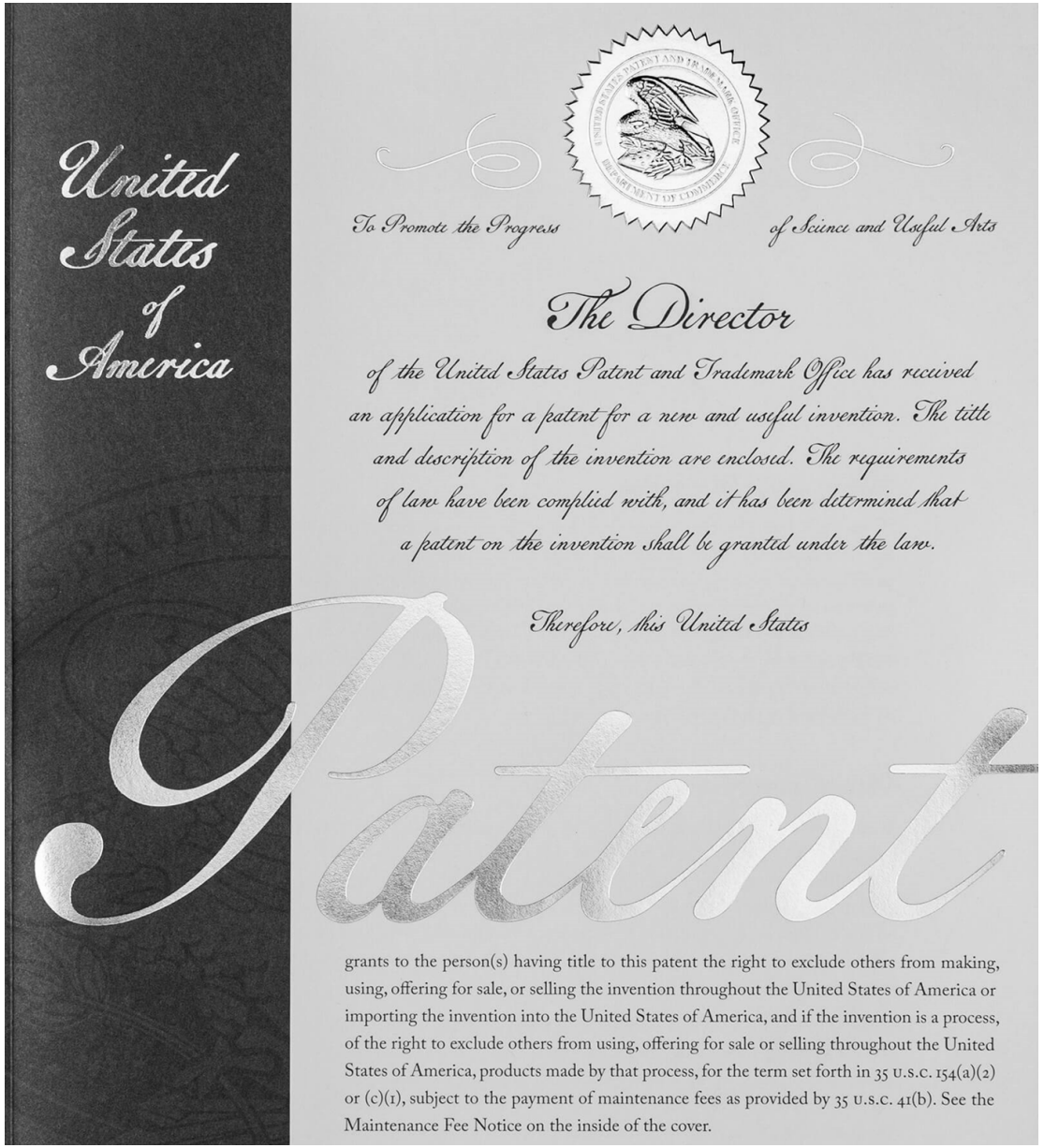
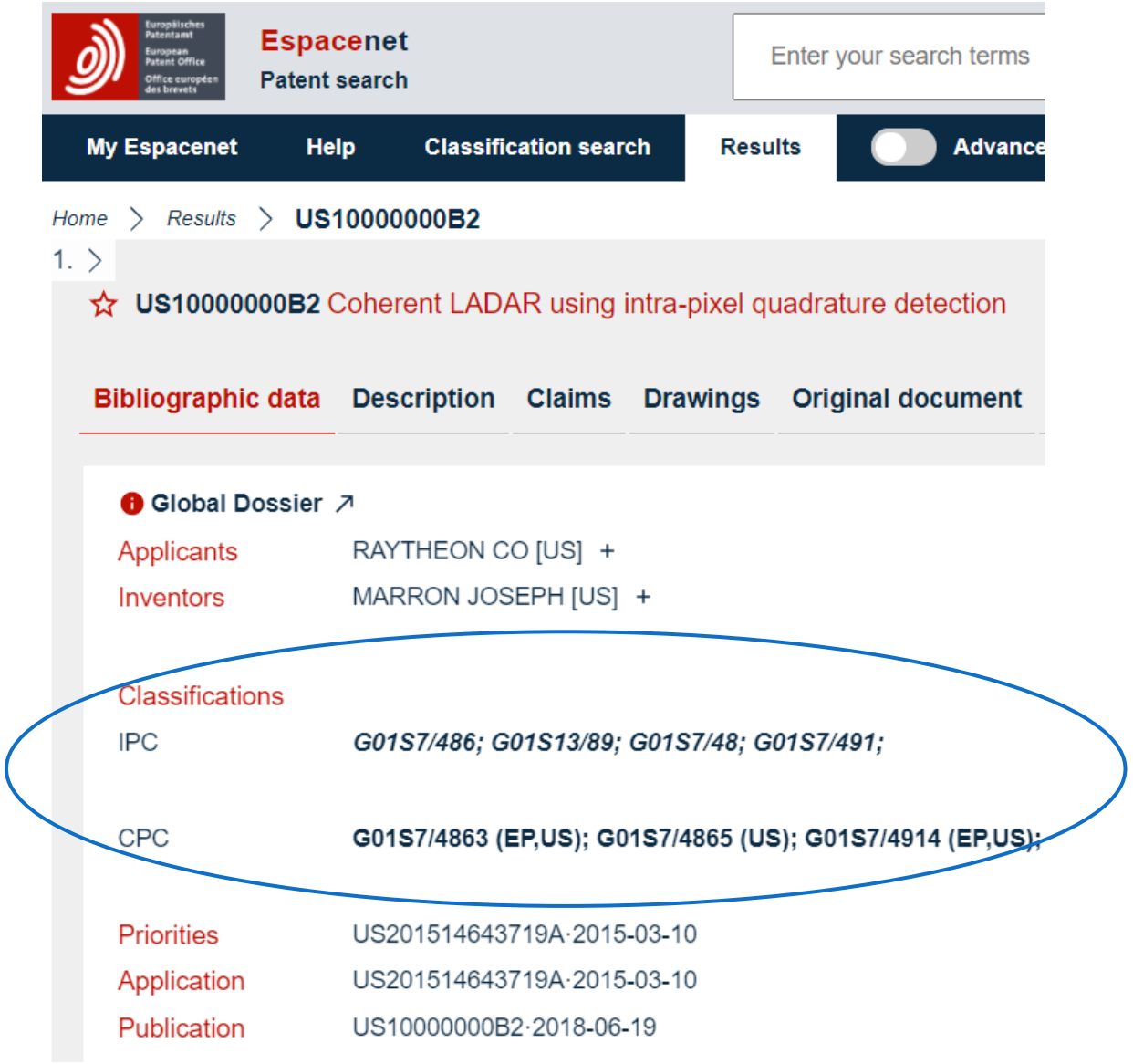
USING NATURAL LANGUAGE PROCESSING TO PREDICT PATENT CLASSIFICATION

CATHERINE FRITZ – CAPSTONE



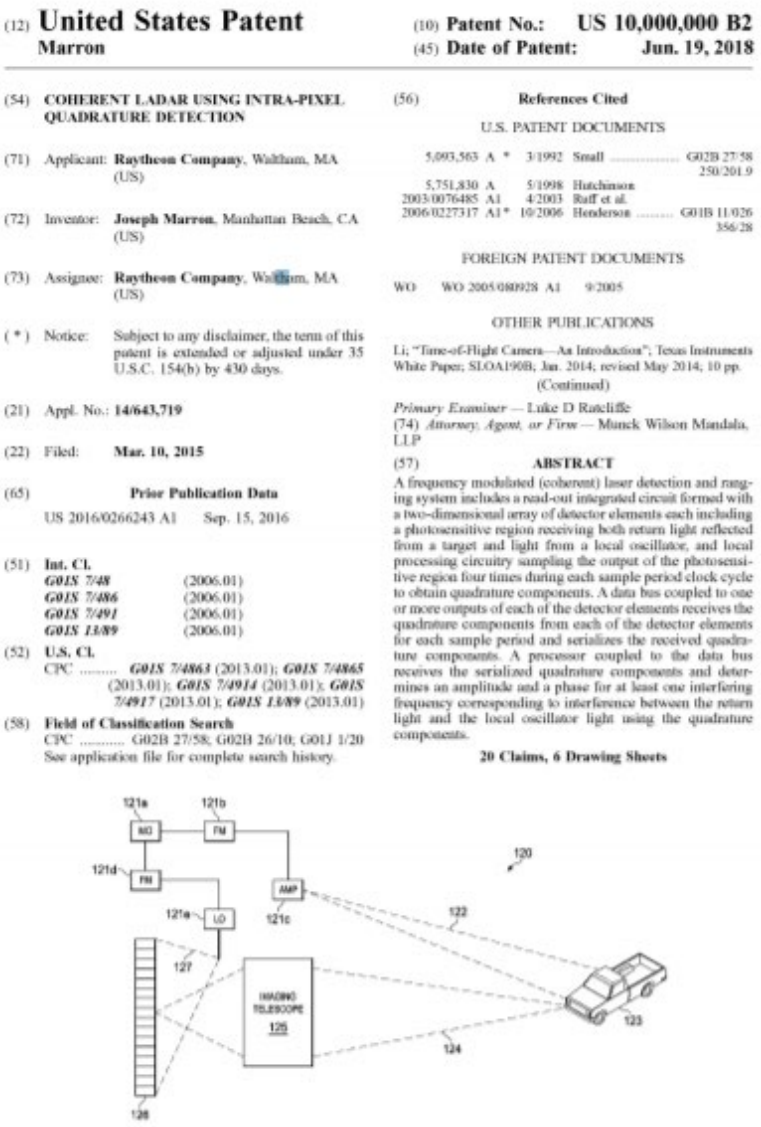
BUSINESS UNDERSTANDING

- Various types of information on patents are connected at various stages in a patent's life, one being the classification of a patent's technology (e.g. mechanical, chemistry, electrical, etc.)
- Classification is assigned by the patent Office some time after a patent is filed
- Could be useful to automate the classification process.



DATA UNDERSTANDING

- 4 main parts to a patent:
  - Coverpage – bibliographic information
  - Specification – technical background
  - Drawings
  - Claims - aspects of an invention that having legal protection
- Use 1st Claim



(12) **United States Patent**
Marron

(10) **Patent No.: US 10,000,000 B2**
(45) **Date of Patent: Jun. 19, 2018**

(54) **COHERENT LADAR USING INTRA-PIXEL QUADRATURE DETECTION**

(71) Applicant: **Raytheon Company**, Waltham, MA (US)

(72) Inventor: **Joseph Marron**, Manhattan Beach, CA (US)

(73) Assignee: **Raytheon Company**, Waltham, MA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 430 days.

(21) Appl. No.: **14/643,719**

(22) Filed: **Mar. 10, 2015**

(65) **Prior Publication Data**

US 2016/0266243 A1 Sep. 15, 2016

(51) **Int. Cl.**

G01S 7/48	(2006.01)
G01S 7/486	(2006.01)
G01S 7/491	(2006.01)
G01S 13/89	(2006.01)

(52) **U.S. Cl.**

CPC ***G01S 7/4863*** (2013.01); ***G01S 7/4865*** (2013.01); ***G01S 7/4914*** (2013.01); ***G01S 7/4917*** (2013.01); ***G01S 13/89*** (2013.01)

(58) **Field of Classification Search**

CPC G02B 27/58; G02B 26/10; G01J 1/20

See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,093,563	A *	3/1992	Small	G02B 27/58 250/201.9
5,751,830	A	5/1998	Hutchinson	
2003/0076485	A1	4/2003	Ruff et al.	
2006/0227317	A1 *	10/2006	Henderson	G01B 11/026 356/28

FOREIGN PATENT DOCUMENTS

WO	WO 2005/080928	A1	9/2005

OTHER PUBLICATIONS

Li, "Time-of-Flight Camera—An Introduction"; Texas Instruments White Paper; SLOA190B; Jan. 2014; revised May 2014; 10 pp.

(Continued)

Primary Examiner — Luke D Ratcliffe
(74) *Attorney, Agent, or Firm* — Munck Wilson Mandala, LLP

(57) **ABSTRACT**

A frequency modulated (coherent) laser detection and ranging system includes a read-out integrated circuit formed with a two-dimensional array of detector elements each including a photosensitive region receiving both return light reflected from a target and light from a local oscillator, and local processing circuitry sampling the output of the photosensitive region four times during each sample period clock cycle to obtain quadrature components. A data bus coupled to one or more outputs of each of the detector elements receives the quadrature components from each of the detector elements for each sample period and serializes the received quadrature components. A processor coupled to the data bus receives the serialized quadrature components and determines an amplitude and a phase for at least one interfering frequency corresponding to interference between the return light and the local oscillator light using the quadrature components.

20 Claims, 6 Drawing Sheets

CLAIM EXAMPLE

- One sentence per claim
- Focused on invention
- Structured around features of invention

What is claimed is:

1. A laser detection and ranging (LADAR) system, comprising:

a two-dimensional array of detector elements, each detector element within the array including:

a photosensitive region configured to receive return light reflected from a target and oscillating local light from a local light source, and

local processing circuitry coupled to an output of the respective photosensitive region and configured to receive an analog signal on the output and to sample the analog signal a plurality of times during each sample period clock cycle to obtain a plurality of components for a sample during each sample period clock cycle;

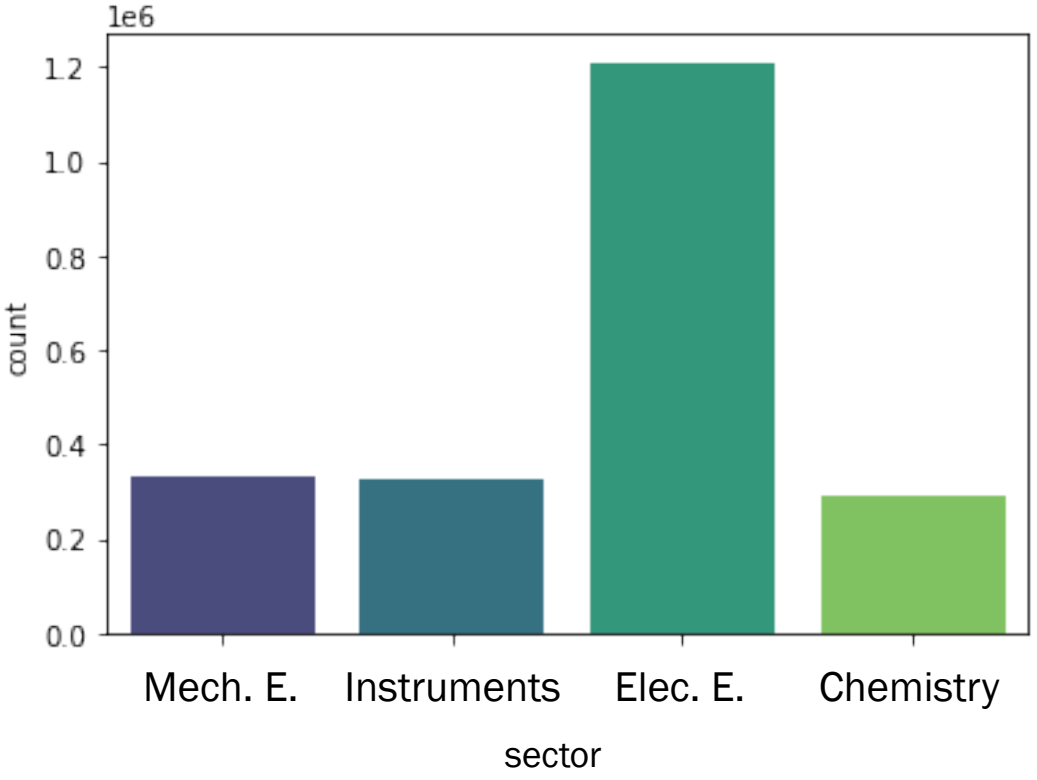
a data bus coupled to one or more outputs of each of the detector elements and configured to receive the plurality of sample components from each of the detector elements for each sample period clock cycle; and

a processor coupled to the data bus and configured to receive, from the data bus, the plurality of sample components from each of the detector elements for each sample period clock cycle and to determine an amplitude and a phase for an interfering frequency corresponding to interference between the return light and the oscillating local light using the plurality of sample components.

2. The system according to claim **1**, wherein the two-dimensional array of detector elements comprises a large format array.

DATA PREPARATION

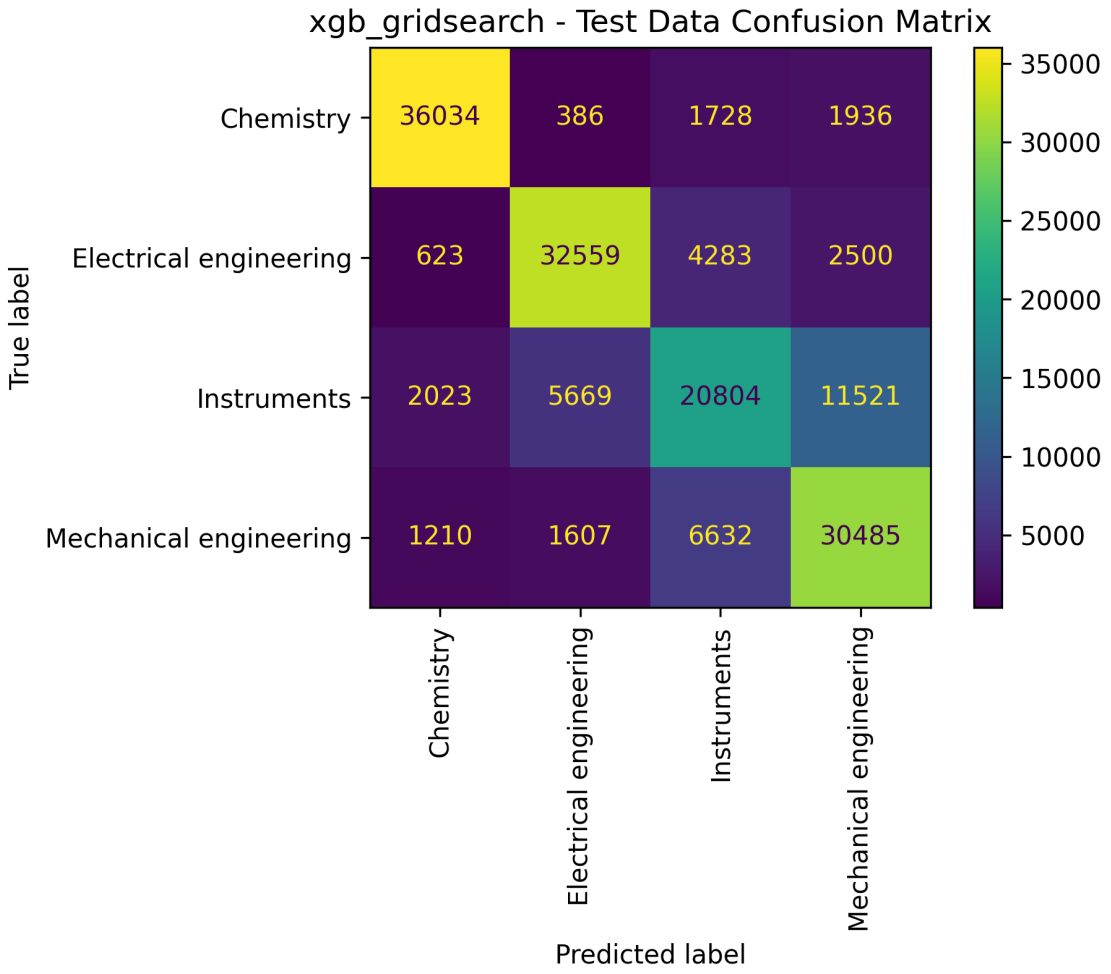
Distribution of Classes for Patents granted since 2011



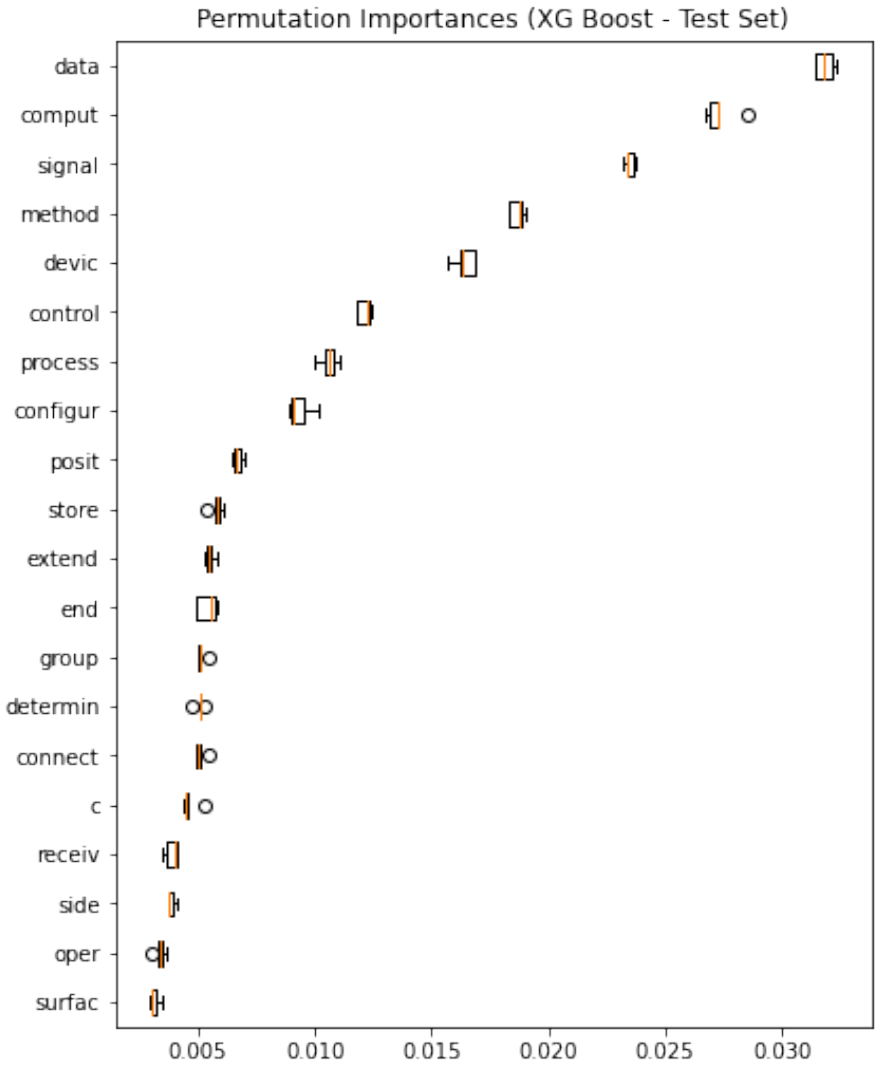
- Classifications:
  - Mechanical engineering
  - Electrical engineering
  - Chemistry
  - Instruments
- Data imbalance - EE

MODEL – XG BOOST

- Best Model: XG Boost
- Captures classifications
- Training Accuracy is: 72.1%
- Validation Accuracy is: 72.0%



EVALUATION



- From these importances, we can see the top 5 words are:
  - data
  - comput (stem of computer, computing, computation, etc.)
  - signal
  - method
  - devic (stem of device, devices, etc.)
- Ramifications for instruments class

RECOMMENDATIONS

- Machine learning could be used to help automate the classification process for patent Offices or third parties.
- Proof of concept for usefulness of machine learning

FUTURE WORK

1. Apply model to more complex classification systems
2. Finalize deep learning model

THANK YOU

NAME: CATHERINE FRITZ

EMAIL: CMFRITZ0@GMAIL.COM

GITHUB: @CMFRITZ

LINKEDIN: LINKEDIN.COM/IN/CATFRITZ

ADDITIONAL INFORMATION CAN BE FOUND AT

HTTPS://GITHUB.COM/CMFRITZ/NLP-PATENT-CLASSIFICATION