# PDT Data Analytics Tool

## (Artificial Intelligence Technology)

## About Me

**Name: Arthur Kang**

**Position:** SW Engineering Intern

**Education:** UC Berkeley

**Major:** Computer Science

Berkeley UNIVERSITY OF CALIFORNIA

## About Team

**Department:** Test Engineering

**Manager:** Anju Gaur

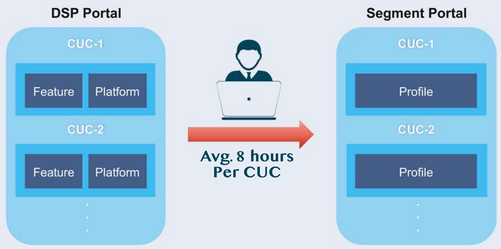
**Collaborators:** Ashok Jude, Harish Gona, Leo Cai

JUNIPER NETWORKS

### Problem Statement

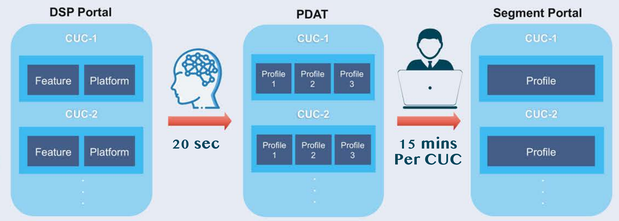
* Currently, when a Customer Use Case (CUC) is submitted, an engineer needs to manually classify this into Profile.

* This process is painstaking and is vulnerable to human errors




* CUC has what Features and Platforms the customer wants to use
* Same Profile = Similar sets of Features and Platforms

### Solution: PDAT (PDT Data Analytics Tool)




* PDAT will find and suggest Top 3 Matching Profiles based on Features and Platforms

* Uses Machine Learning, Neural Network, and Artificial Intelligence technology

### Data

| | ucId | featuresUsed | products | Profile |
|---|---|---|---|---|
| 1565 | | ECMP,L2 Circuit,LDP Track-IGP,LDP igp-synchronization,LDP over RSVP,LSP - Link Protection,MPLS,MPLS FRR - Link Protection,MPLS FRR - Node Protection,OSPF V1/V2,SSH,SSH v2,RSVP (Resource Reservation Protocol),RSVP TE Admin Groups | QFX5100 | ENTDC-Edge-VPNTunnel |
| 1457 | | 802.1Q,ACL,AE Bundle,Aggregate Ethernet,BFD,BGP Community,BGP EBGP Multihop,BGP as-path,BGP flow spec,BGP inet-label-unicast,BGP v4,BGP v6,BGP-LS,BGP-LU,BMPv3,CoS,CoS (Class of service),CoS BA Classifier,CoS Rewrite,CoS Scheduler,CoS VoQ,ECMP,GRE,ICMP,IPV6,IPv4 static routing,IPv6 static routing,ISIS,ISIS LFA (For 17.2),LACP,LDP (MPLS Label Distribution Protocol),LDP Track-IGP,LDP igp-synchronization,LFA ISIS,LLDP,MBGP,MPLS,MPLS FRR - Link Protection,MPLS FRR - Node Protection,Micro BFD,MicroBFD,Multicast,NTP,Netconf,PIM | QFX10003 | ENTDC-Edge-VPNTunnelMulticastL3L2 |

### Performance


### PDAT Output

| | Prediction | Answer | Probability |
|---|---|---|---|
| 27 | SP-MPLS-VPNServices | SP-MPLS-VPNServices | [(SP-MPLS-VPNServices, 0.9832184), (SP-IP-BGPSERVICES, 0.0054024328), (SP-METROSERVICES, 0.0032732305)] |
| 28 | DCPublicCF-IPCLOS-IBGP,Aã | DCPublicCF-IPCLOS-IBGP,Aã | [(DCPublicCF-IPCLOS-IBGP,Aã, 0.9999584), (ENTCnB-AccessAggregation-Legacy, 1.2084683e-05), (SP-MPLS-VPNServices, 6.294559e-06)] |
| 29 | ENTCnB-AccessAggregation-XSTP | SP-MPLS-VPNServices | [(ENTCnB-AccessAggregation-XSTP, 0.5934151), (SP-MPLS-VPNServices, 0.16084856), (ENTCnB-AccessAggregation-Legacy, 0.07525084)] |
| 30 | ENTCnB-JFEInterconnect-EVPN-MPLS | ENTCnB-JFEInterconnect-EVPN-MPLS | [(ENTCnB-JFEInterconnect-EVPN-MPLS, 0.9218795), (SP-MPLS-VPNServices, 0.07770566), (CS2, 8.836093e-05)] |
| 31 | ENTCnB-JFEInterconnect-EVPN-MPLS | ENTCnB-JFEInterconnect-EVPN-MPLS | [(ENTCnB-JFEInterconnect-EVPN-MPLS, 0.99973065), (SP-MPLS-VPNServices, 8.172085e-05), (CS2, 4.7590856e-05)] |

## Results

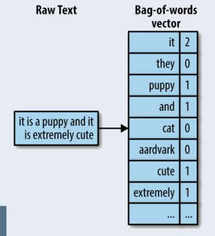
73.3% (Correct Profile for Top 1)
85.5% (Correct Profile Among Top 2)
90.8% (Correct Profile Among Top 3)

Moved to Production
Will be deployed in PDT Dashboard

## Algorithm

### Bag of Words Algorithm

* Counts the number of times each words appear
* Stores the result in a vector




### TF-IDF

* Term Frequency-Inverse Document Frequency
* Converts the word count to word frequency
* Removes words that appear too few/much times


### Keras Neural Network Model

* Analyzes the data with logic structure like how humans draw conclusions
* Creates Artificial Neurons Mimics how human brains work