Auditors Aid

Anomaly detection in
Not-for-Profit data

GOAL

Augmenting the current IRS system-of-review for Not-For-Profit organizations

Use Machine Learning to:

- Minimizing excess paperwork
- Finding Anomalies within submitted forms
- Help concentrating area of focus for Auditors

NTEE Category	Definition
A	Arts, Culture, and Humanities
B	Education
C	Environmental Quality, Protection, and Beautification
D	Animal-Related
E	Health
F	Mental Health, Crisis Intervention
G	Diseases, Disorders, Medical Disciplines
H	Medical Research
I	Crime, Legal Related
J	Employment, Job Related
K	Food, Agriculture, and Nutrition
L	Housing, Shelter
M	Public Safety
N	Recreation, Sports, Leisure, Athletics
O	Youth Development
P	Human Services - Multipurpose and Other
Q	International, Foreign Affairs, and National Security

Each Not-for-Profit has an area of focus, as defined by the National Taxonomy of Exempt Entities (NTEE) Classification System

Each group has a unique footprint

R	Civil Rights, Social Action, Advocacy
S	Community Improvement, Capacity Building
T	Philanthropy, Voluntarism, and Grantmaking Foundations
U	Science and Technology Research Institutes, Services
V	Social Science Research Institutes, Services
W	Public, Society Benefit - Multipurpose and Other
X	Religion Related, Spiritual Development
Y	Mutual/Membership Benefit Organizations, Other
Z	Unknown

DATA

IRS Exempt Organizations Business Master File Extract (EO BMF)

Consists of 1,691,200 Not-for-Profit

Each Not-for-Profit has 28 self-reported characteristics

	EIN	STREET	CITY	STATE	ZIP	GROUP	SUBSECTION	AFFILIATION	CLASSIFICATION	DEDUCTIBILITY	...	STATUS	FILING_REQ_CD	PF_FILING_REQ_CD	ACCT_PD	ASSET_AMT
12	10002847	PO BOX 126	HULLS COVE	ME	04644-0126	0	3	3	1000	1	...	1	2	0	12	0.0
13	10011694	30 SAMUEL PRESCOTT DR	STOW	MA	01775-1167	0	3	3	1000	1	...	1	2	0	12	0.0
14	10015091	PO BOX 13	CEDAR KNOLLS	NJ	07927-0013	0	3	3	1000	1	...	1	1	0	12	62462.0
15	10017496	PO BOX 534	YORK HARBOR	ME	03911-0534	0	3	3	2000	1	...	1	1	0	12	233819.0
16	10018555	81 COLLEGE AVE	ORONO	ME	04473-4210	287	7	9	1000	2	...	1	2	0	12	0.0

Form **990**

Department of the Treasury
Internal Revenue Service

Return of Organization Exempt From Income Tax

Under section 501(c), 527, or 4947(a)(1) of the Internal Revenue Code (except private foundations)

▶ Do not enter social security numbers on this form as it may be made public.

▶ Go to *www.irs.gov/Form990* for instructions and the latest information.

OMB No. 1545-0047

20**18**

Open to Public Inspection

A For the 2018 calendar year, or tax year beginning _____, 2018, and ending _____, 20___

B Check if applicable:	**C** Name of organization	**D** Employer identification number
☐ Address change	Doing business as	
☐ Name change	Number and street (or P.O. box if mail is not delivered to street address) Room/suite	**E** Telephone number
☐ Initial return		
☐ Final return/terminated	City or town, state or province, country, and ZIP or foreign postal code	
☐ Amended return		**G** Gross receipts $
☐ Application pending	**F** Name and address of principal officer:	**H(a)** Is this a group return for subordinates? ☐ Yes ☐ No
		H(b) Are all subordinates included? ☐ Yes ☐ No
		If "No," attach a list. (see instructions)

I Tax-exempt status: ☐ 501(c)(3) ☐ 501(c) (   ) ◄ (insert no.) ☐ 4947(a)(1) or ☐ 527

J Website: ▶

H(c) Group exemption number ▶

K Form of organization: ☐ Corporation ☐ Trust ☐ Association ☐ Other ▶ _____   **L** Year of formation: _____   **M** State of legal domicile: _____

Part I   Summary

Activities & Governance

1 Briefly describe the organization's mission or most significant activities: _____

_____

2 Check this box ▶ ☐ if the organization discontinued its operations or disposed of more than 25% of its net assets.

3 Number of voting members of the governing body (Part VI, line 1a)	**3**	
4 Number of independent voting members of the governing body (Part VI, line 1b)	**4**	
5 Total number of individuals employed in calendar year 2018 (Part V, line 2a)	**5**	
6 Total number of volunteers (estimate if necessary)	**6**	
7a Total unrelated business revenue from Part VIII, column (C), line 12	**7a**	
b Net unrelated business taxable income from Form 990-T, line 38	**7b**	

		Prior Year	Current Year
8	Contributions and grants (Part VIII, line 1h)		

DATA - Contents

Column Name	Contents
EIN	Employer Identification Number (EIN)
NAME	Primary Name of Organization
ICO	In Care of Name
STREET	Street Address
CITY	City
STATE	State
ZIP	Zip Code
GROUP	Group Exemption Number
SUBSECTION	Subsection Code
AFFILIATION	Affiliation Code(s)
CLASSIFICATION	Classification Code(s)
RULING	Ruling Date
DEDUCTIBILITY	Deductibility Code
FOUNDATION	Foundation Code

Column Name	Contents
ACTIVITY	Activity Codes
ORGANIZATION	Organization Code
STATUS	Exempt Organization Status Code
TAX_PERIOD	Tax Period
ASSET_CD	Asset Code
INCOME_CD	Income Code
FILING_REQ_CD	Filing Requirement Code
ACCT_PD	Accounting Period
ASSET_AMT	Asset Amount
INCOME_AMT	Income Amount (includes negative sign if amount is negative)
REVENUE_AMT	Form 990 Revenue Amount (includes negative sign if amount is negative)
NTEE_CD National	Taxonomy of Exempt Entities (NTEE) Code
SORT_NAME	Sort Name (Secondary Name Line)

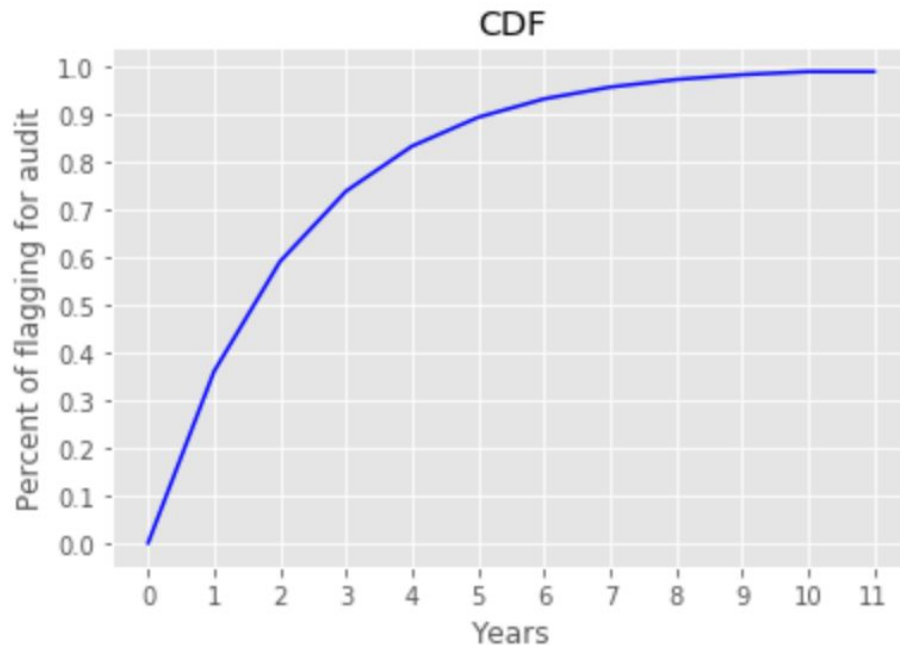# Results for predicting Major Class

My Result:

36%

True Random:

4%

Results for predicting Major Class

Given a misclassification exists,

There is 60% probability it will be detected by the 2nd year of reporting, reaching 95% at the 7th year.

CDF

Benefits

- Finding mistakes quicker and reducing paperwork

- Automatically identifying errors on submitted forms

- Flagging area of concentrating for Auditors

Next Steps

Link other tax form documentation to prediction criteria

Refine NTEE criteria to fully utilize entire data set

Find opportunities to incorporate and test out findings