

Bauhaus University Weimar
Faculty of Media
Computer Science for Digital Media

Detecting Web Functions

Master's Thesis Defense

Submitted by Ademola Eric Adewumi

20.10.2021

Overview

Major sections of this thesis defense presentation

Introduction

Objectives

Web Page Functions

Dataset

Annotation

Classification

Prediction and Conclusion

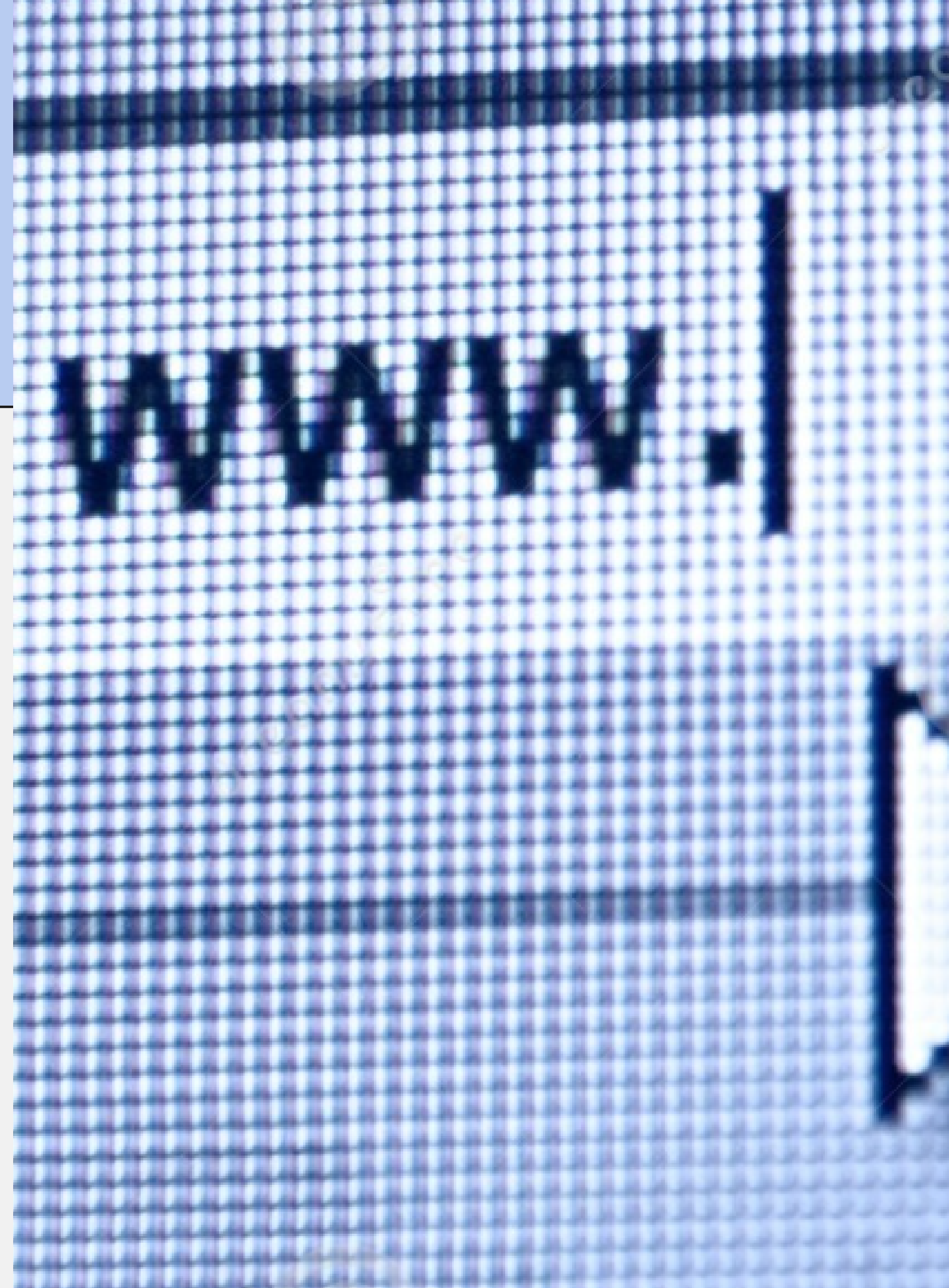
Introduction

The WWW is made up of HTML webpages

Each Webpage reflects authors intentions

Similar webpages look alike

We want to train machines to identify webpages like humans do



Objective

Build a Classifier

Using the largest publicly available
annotated dataset

Web Page Functions

Entertainment

Netflix, Amazon Prime
Video

Data Collection

Google Forms

eCommerce

Alibaba, ebay

Discussion

Discord, Yahoo Messenger

Information

CNN, BBC

Link/ Suggestion

Nairaland

Task 1

[Bossier City](#)

[About Bossier City](#)
[Hotels](#)
[Holiday Rentals](#)
[Restaurants](#)
[Things to do](#)
[Flights](#)

[United States](#) > [Louisiana \(LA\)](#) > [Bossier City](#) > [Bossier City Hotels](#)

La Quinta Inn Bossier City

376 Reviews
 309 Preston Blvd., Bossier City, LA 71201

Enter your travel dates

Nov. 2017							Dec. 2017						
S	M	T	W	T	F	S	S	M	T	W	T	F	S
1	2	3	4	5	6	7	1	2	3	4	5	6	7
8	9	10	11	12	13	14	8	9	10	11	12	13	14
15	16	17	18	19	20	21	15	16	17	18	19	20	21
22	23	24	25	26	27	28	22	23	24	25	26	27	28
29	30						29	30					

Lowest prices found

1 room • 2 adults

- Save \$13 with Expedia Opodo **EUR 73*** [View Deal](#)
- Save \$18 with Expedia Orbitz **EUR 74*** [View Deal](#)

*Prices are provided by our partners, and reflect taxes.

All photos (92)

[Overview](#)
[Rooms](#)
[Reviews](#)
[About](#)
[Photos](#)
[Neighborhood](#)
[Q&A](#)
[Room Tips](#)

Overview

3.5
 376 reviews

Category	Percentage
Excellent	27%
Very good	31%
Average	16%
Poor	13%
Terrible	13%

Free WiFi
Free Parking
Breakfast included
Pool
Non-Smoking Rooms

Suites
Pets Allowed
★ ★ ★ ★ ☆
2.5 Star Hotel
[Add hotel details](#)

TRAVELLERS TALK ABOUT

- "room was clean" (29 reviews)
- "quiet area" (4 reviews)
- "older hotel" (4 reviews)

OFFERS FROM LA QUINTA INN BOSSIER CITY

- Special Offer: Save up to 30% [View all deals](#)
- Hotel deals

Map data ©2017 Google

Rooms

17 Nov. – 18 Nov. • | 1 room, 2 adults, 0 children •

Standard Room

Free Cancellation

Select

Show Available Rooms •

Reviews (376)

[Write a Review](#)

Task 3

[Markets](#)
[Economy](#)
[Companies](#)
[Tech](#)
[Autos](#)
[India](#)
[Video](#)

Mitsubishi Corp

46.26

\$ -0.75 / -1.60%

+7.98%

Quote

Profile

News

Charts

Forecasts

Financials

Shareholders

Competitors

1 Year

In Distribution Services

News

Real Estate

PREVIOUS CLOSE: \$47.00
\$ DAILY CHANGE: +0.80%

Date	Price	% Change
Jan 2015	\$40.10	-0.41%
May 2016	\$45.00	+1.47%
Nov 2016	\$45.00	-0.00%
Dec 2016	\$45.00	-0.00%
Nov 2017	\$46.26	+0.28%

Latest MSBHY News | Press Releases

No recent news for Mitsubishi.

Today's Trading

Growth & Valuation

Previous close	\$47.00	Earnings growth (last year)	--
Today's open	\$46.30	Earnings growth (this year)	+18.60%
Day's range	\$46.26 - \$46.30	Earnings growth (next 5 years)	+4.40%
Volume	3,740	Revenue growth (last year)	+0.80%
Average volume (3 months)	6,238	P/E ratio	--
Market cap	\$27.3B	Price/Sales	0.96
Dividend yield	3.60%	Price/Book	0.94

Data as of 5:00pm ET, 11/21/2017

Competitors

Financials

No competitors data available.

Next reporting date	November 5, 2017
EPS forecast (this quarter)	\$1.32
Annual revenue (last year)	\$10.3B
Annual profit (last year)	\$4.1B
Net profit margin	0.60%

Profile

Forecasts

Sector

Distribution Services

Industry

Wholesale Distributors

President & Representative Director

Takafusa Katsuhiko

Operations Manager & Executive Officer

Mitsuyuki Takada

Corporate Headquarters

Tokyo, Tokyo

HISTORICAL FORECAST

Year	EPS	% Change
2015	\$1.10	+0.00%
2016	\$1.20	+9.09%
2017	\$1.30	+8.33%

Last 12 Months Next 12 Months

The Motley Fool

Fast Follower

This Stock Could Be Like Buying Amazon for \$3.19

37-Year-Old CEO Bets \$782,000,000 on 1 Stock

You Can Still Buy This "Millionaire Maker" Stock

Apple Quietly Leases 5,000 Acre Abandoned Military Base

Motley Fool Issues Rare Triple-Buy Alert

Partner Offers

SmartAsset

German Election, How will markets react? - capital at risk

Hire a pro: compare top 7 financial advisors in your area

Dataset

8,940

Webpages

Derived from crowd sourced Annotations

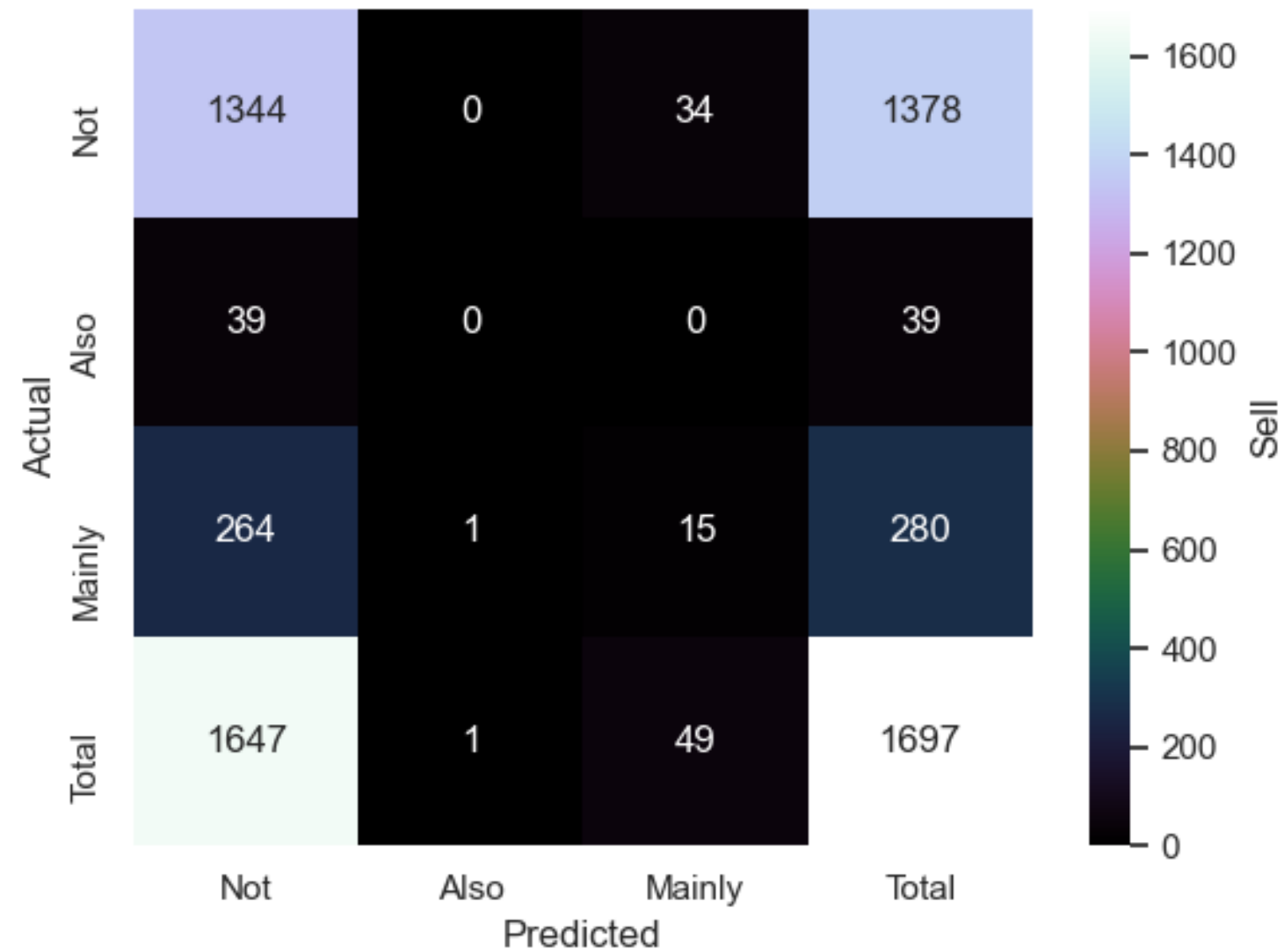
Annotations

Amazon Mechanical
Turk

Multi Annotator Competency
Estimator

Accept/Reject	Worker	Sec.	Check	Mainly	Task 1	Discuss	Entert.	Form	Inform	Link	Sell	Comm.
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	A1W..T46 (0.25)	57		inform		not	mainly	not	not	not	not	
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	A2F..CRS (0.44)	186		entertain		also	mainly	not	also	also	not	
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	APF..HBC (0.56)	237		inform		not	mainly	not	also	not	not	
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	A29..8MS (0.66)	189		inform		not	also	mainly	also	also	not	
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	AUN..UO2 (0.36)	76		inform		not	mainly	not	not	not	not	
Accept/Reject	Worker	Sec.	Check	Mainly	Task 1	Discuss	Entert.	Form	Inform	Link	Sell	Comm.
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	A1W..T46 (0.25)	56		entertain		not	not	not	mainly	not	not	
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	APF..HBC (0.56)	266		entertain		not	also	mainly	also	not	not	
Accept/Reject	Worker	Sec.	Check	Mainly	Task 1	Discuss	Entert.	Form	Inform	Link	Sell	Comm.
<div>Approve</div> <div>Reject internally</div> <div>Reject</div>	A1W..T46 (0.25)	89		form		not	not	not	mainly	not	not	

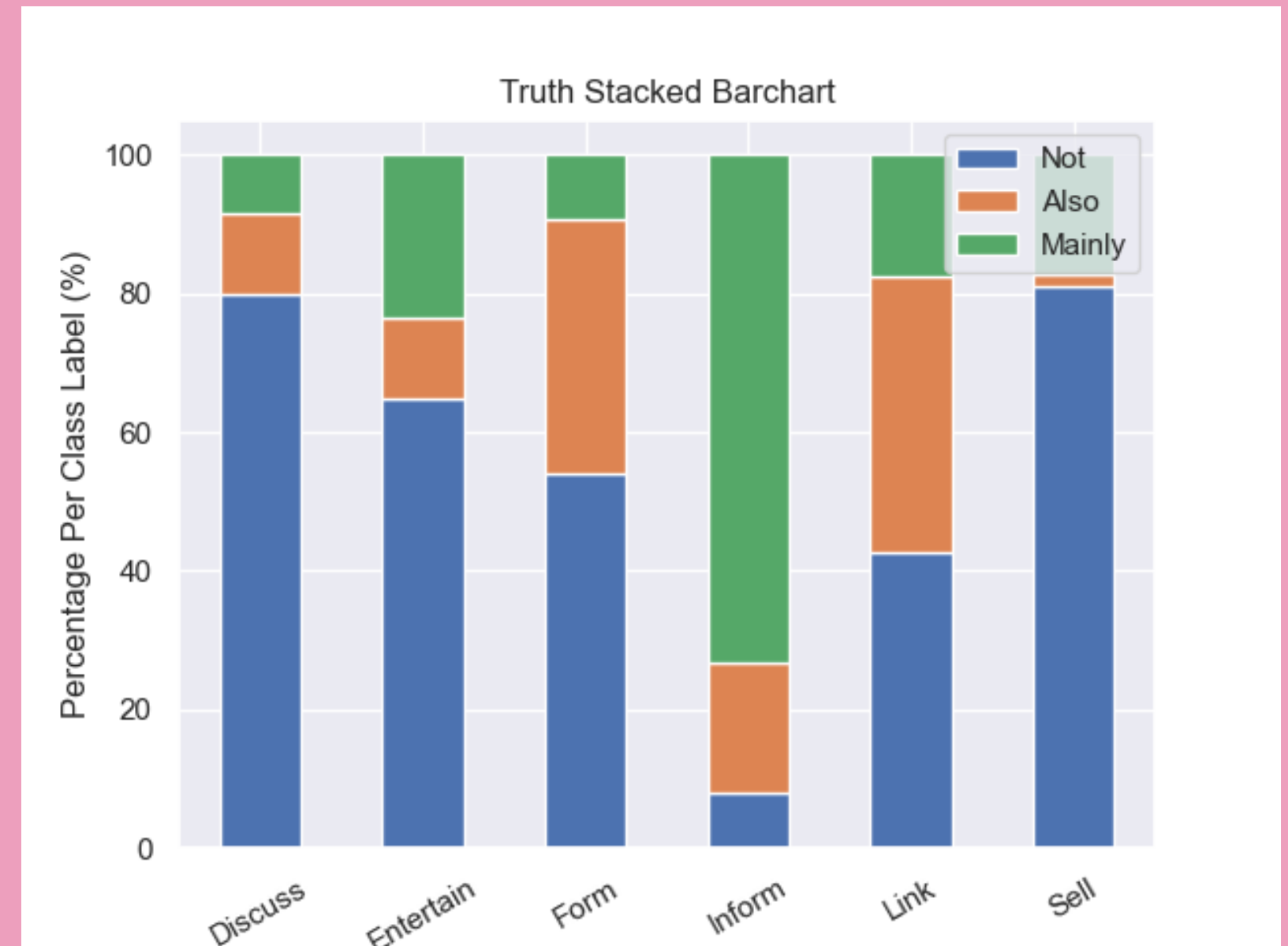
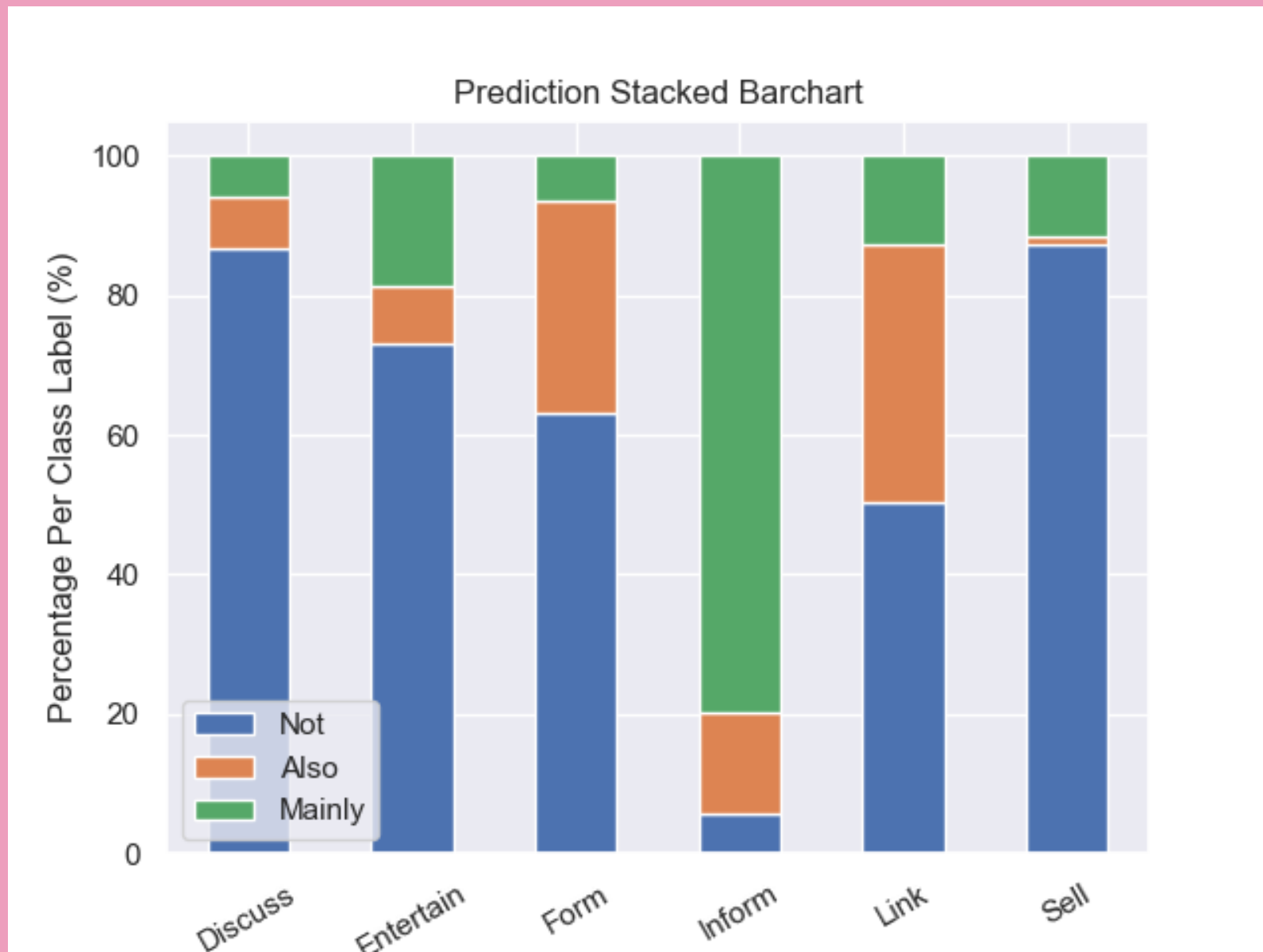
Classification



Random Forest Classifier

We used Random Forest classification because it is a forest of decision trees with oblique hyperplanes

Results



THANK YOU