



Contextual Ads

as the New Frontier in Privacy Compliant Advertising

Contents



Current digital advertising landscape - what's at stake?



How can Machine Learning in contextual advertising contribute to addressing privacy concerns?



What is the impact in terms of dollars ?



What's our data - how does it look like?



Preliminary results from EDA



Next steps...

Current digital advertising landscape - what's at stake?

The screenshot shows the Icelandair website interface with several browser overlays. On the left, the 'About icelandair.com' panel displays site security and permissions. In the center, a 'Cookies in use' panel lists cookies from icelandair.com and smaato.net, with details for a specific NID cookie. The background shows the flight search form with fields for 'From' and 'To'. A blue text overlay reads 'but... what is a cookie?'. On the right, an illustration of a chocolate chip cookie with crumbs is shown. The bottom navigation bar includes links for 'Book flights', 'Check-in', and 'Manage booking'.

Flights to Europe & Iceland | Icel. x +

https://www.icelandair.com/en-ca/

About icelandair.com

- Connection is secure
- Permissions for this site
- Cookies (95 cookies in use)
- Tracking prevention for this site (Balanced)
- Trackers (37 blocked)

Flights Vacations Cars

Your next adventure is just a search away

From
Departing from

To
Where are you going?

but... what is a cookie?

Sign up!

Cookies in use

Allowed Blocked

The following cookies were set when you viewed this page

- NID
- icelandair.com
- smaato.net

Name	NID
Content	511=Udywz2Z5tDKJNsT95Gxj_9au4xvllIH!wxul6VgAj...
Domain	.google.com
Path	/
Send for	Secure connections only
Created	Thursday, June 22, 2023 at 2:35:59 PM
Expires	Friday, December 22, 2023 at 1:35:59 PM

Block Remove Done

Book flights Check-in Manage booking

How can Machine Learning in contextual advertising contribute to addressing privacy concerns?

Propo

The screenshot shows the Guardian website interface. At the top, there is a travel advertisement for Icelandair with the text "Flight sale to Europe!" and "From \$799". Below the ad, there is a navigation bar with links for "Support the Guardian", "Print subscriptions", "Sign in", "Search jobs", "Search", and "International edition". The main content area features a news article titled "Lancashire holidays" by Chris Moss, dated Thu 29 Jun 2023 07:00 BST. The article includes a photograph of a river scene and text about the region's old mills and museums.



Picks advertisers
(IcelandAir) who wants
to show travel ad

of a cookie (tracking users across web) to show an ad?



Ad platform receives
the category signal
Example: Travel



ML Predicts the
category it belongs



What is the impact in terms of dollars ?

“..the Interactive Advertising Bureau estimates publishers' exposure at up to **\$10 billion in lost revenue** without third party **cookies**..” Source: [IAB](#)

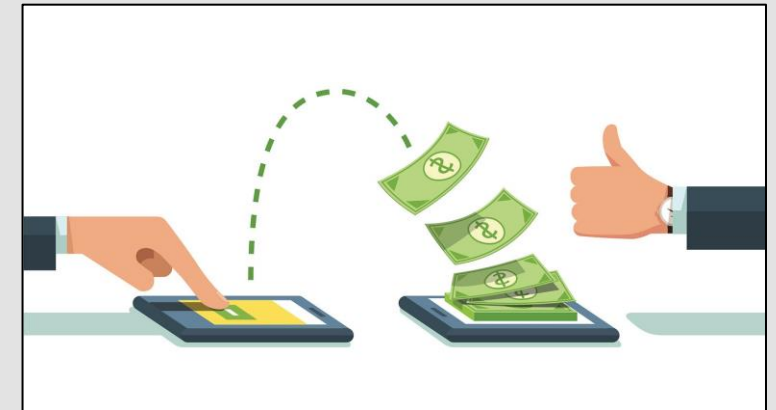
Impact of our proposed solution



Users seeing relevant ads



Advertisers will gain high ROI



Publishers sell more ads

What's our data - how does it look like?



This dataset contains around 210,000 news headlines from 2012 to 2022 from HuffPost.

209,527 rows
&
6 columns

link	headline	category	short_description	authors	date
https://www.huffpost.com/entry/covid-boosters-...	Over 4 Million Americans Roll Up Sleeves For O...	U.S. NEWS	Health experts said it is too early to predict...	Carla K. Johnson, AP	2022-09-23
https://www.huffpost.com/entry/american-airlin...	American Airlines Flyer Charged, Banned For Li...	U.S. NEWS	He was subdued by passengers and crew when he ...	Mary Papenfuss	2022-09-23
https://www.huffpost.com/entry/funniest-tweets...	23 Of The Funniest Tweets About Cats And Dogs ...	COMEDY	"Until you have a dog you don't understand wha...	Elyse Wanshel	2022-09-23
https://www.huffpost.com/entry/funniest-parent...	The Funniest Tweets From Parents This Week (Se...	PARENTING	"Accidentally put grown-up toothpaste on my to...	Caroline Bologna	2022-09-23
https://www.huffpost.com/entry/amy-cooper-lose...	Woman Who Called Cops On Black Bird-Watcher Lo...	U.S. NEWS	Amy Cooper accused investment firm Franklin Te...	Nina Golgowski	2022-09-22
https://www.huffpost.com/entry/belk-worker-fou...	Cleaner Was Dead In Belk Bathroom For 4 Days B...	U.S. NEWS	The 63-year-old woman was seen working at the ...		2022-09-22
https://www.huffpost.com/entry/reporter-gets-a...	Reporter Gets Adorable Surprise From Her Boyfr...	U.S. NEWS	"Who's that behind you?" an anchor for New Yor...	Elyse Wanshel	2022-09-22
https://www.huffpost.com/entry/puerto-rico-wat...	Puerto Ricans Desperate For Water After Hurric...	WORLD NEWS	More than half a million people remained witho...	DÁNICA COTO, AP	2022-09-22
https://www.huffpost.com/entry/mija-documentar...	How A New Documentary Captures The Complexity ...	CULTURE & ARTS	In "Mija," director Isabel Castro combined mus...	Marina Fang	2022-09-22
https://www.huffpost.com/entry/biden-un-russia...	Biden At UN To Call Russian War An Affront To ...	WORLD NEWS	White House officials say the crux of the pres...	Aamer Madhani, AP	2022-09-21

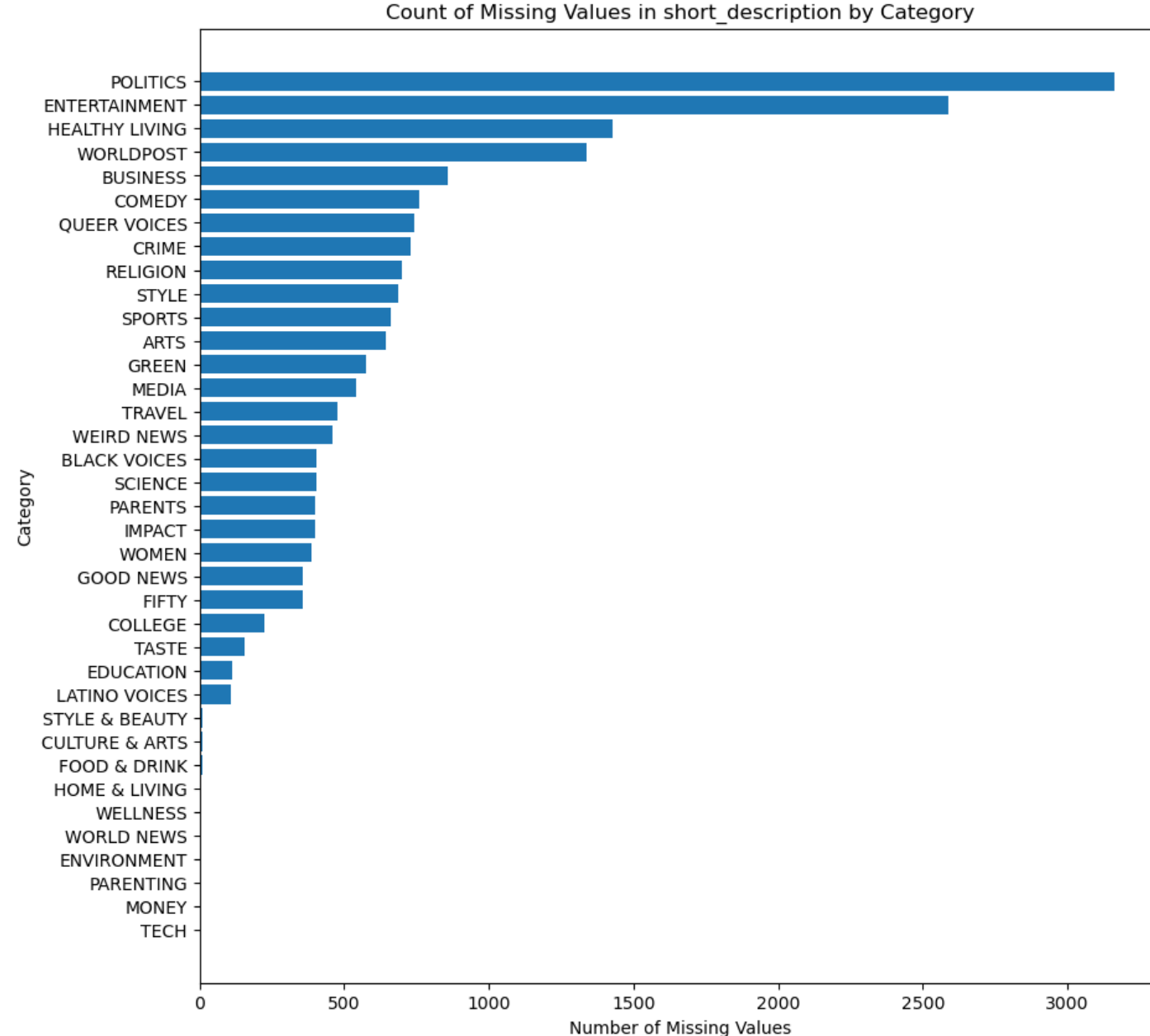
Preliminary results from EDA

Number of missing values in each variable:

link	0
headline	6
category	0
short_description	19712
authors	37418
date	0
month_num	0
year	0

Percentage missing values in each variable:

link	0.00
headline	0.00
category	0.00
short_description	9.41
authors	17.86
date	0.00
month_num	0.00
year	0.00



Next steps...

- Advanced EDA

Example: Understanding top words by category, trends by month/year etc..

- Grouping similar categories.

Example: “Science” , “Tech” -> ‘Science & Tech’

- Data Preparation:

- (i) Tokenizing, stemming and lemmatization
- (ii) Vectorization

- Exploring Multi-classification Models



Thank You!