

# ShritejShrikant\_Chavan\_HW8

December 16, 2023

## 0.1 Data Preparation

### 0.1.1 Import Packages

```
[1]: # check if we are using google colab
from pathlib import Path
import textwrap
if 'google.colab' in str(get_ipython()):
    from google.colab import drive
    drive.mount("/content/drive")
    !pip install datasets transformers bertopic umap hdbscan openai -U -qq

    base_folder = Path("/content/drive/MyDrive/NLP")
else:
    base_folder = Path("/home/harpreet/Insync/google_drive_shaannoor/data")
```

Mounted at /content/drive

521.2/521.2

kB 7.5 MB/s eta 0:00:00

8.2/8.2 MB

39.1 MB/s eta 0:00:00

154.1/154.1

kB 13.0 MB/s eta 0:00:00

Preparing metadata (setup.py) ... done

5.2/5.2 MB

39.6 MB/s eta 0:00:00

Installing build dependencies ... done

Getting requirements to build wheel ... done

Preparing metadata (pyproject.toml) ... done

221.4/221.4

kB 28.7 MB/s eta 0:00:00

115.3/115.3

kB 16.6 MB/s eta 0:00:00

134.8/134.8

kB 20.5 MB/s eta 0:00:00

90.9/90.9 kB

```

11.6 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
                        86.0/86.0 kB
10.8 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
                        75.0/75.0 kB
9.4 MB/s eta 0:00:00
                        76.9/76.9 kB
10.0 MB/s eta 0:00:00
                        58.3/58.3 kB
6.9 MB/s eta 0:00:00
                        1.3/1.3 MB
60.1 MB/s eta 0:00:00
                        55.8/55.8 kB
7.1 MB/s eta 0:00:00
  Building wheel for umap (setup.py) ... done
  Building wheel for hdbscan (pyproject.toml) ... done
  Building wheel for sentence-transformers (setup.py) ... done
  Building wheel for umap-learn (setup.py) ... done
ERROR: pip's dependency resolver does not currently take into account all
the packages that are installed. This behaviour is the source of the following
dependency conflicts.

llmx 0.0.15a0 requires cohere, which is not installed.
llmx 0.0.15a0 requires tiktoken, which is not installed.

```

```

[2]: def print_wrap(text, width=80):
      """
      Prints the given text, wrapping lines to a maximum of the specified width_
      ↪(default is 80 characters).

      Args:
      text (str): The text to be printed.
      width (int): The maximum width of a line, in characters.
      """
      wrapper = textwrap.TextWrapper(width=width)
      wrapped_text = wrapper.fill(text)
      print(wrapped_text)

```

```

[3]: data_folder = base_folder/'datasets/hw8'
      model_folder = base_folder/'models/nlp_fall_2023/hw8'
      model_folder.mkdir(exist_ok=True, parents=True)
      data_folder.mkdir(exist_ok=True, parents=True)

```

### 0.1.2 Load Dataset

```
[4]: from datasets import load_dataset
```

```
medium_dataset = load_dataset('csv', data_files= str(data_folder /  
↳ 'titles_cleaned.csv'))
```

Downloading data files: 0%| | 0/1 [00:00<?, ?it/s]

Extracting data files: 0%| | 0/1 [00:00<?, ?it/s]

Generating train split: 0 examples [00:00, ? examples/s]

```
[5]: medium_dataset
```

```
[5]: DatasetDict({  
    train: Dataset({  
        features: ['Unnamed: 0', 'url', 'title', 'claps', 'responses',  
        'readTime', '$oid', '$date', 'datetime', 'year', 'month', 'date', 'hour',  
        'minute', 'second'],  
        num_rows: 28383  
    })  
})
```

```
[6]: medium_dataset['train'].features
```

```
[6]: {'Unnamed: 0': Value(dtype='int64', id=None),  
    'url': Value(dtype='string', id=None),  
    'title': Value(dtype='string', id=None),  
    'claps': Value(dtype='float64', id=None),  
    'responses': Value(dtype='float64', id=None),  
    'readTime': Value(dtype='string', id=None),  
    '$oid': Value(dtype='string', id=None),  
    '$date': Value(dtype='string', id=None),  
    'datetime': Value(dtype='string', id=None),  
    'year': Value(dtype='float64', id=None),  
    'month': Value(dtype='float64', id=None),  
    'date': Value(dtype='float64', id=None),  
    'hour': Value(dtype='float64', id=None),  
    'minute': Value(dtype='float64', id=None),  
    'second': Value(dtype='float64', id=None)}
```

```
[7]: medium_titles = medium_dataset['train']['title']
```

## 0.2 Sentence BERT for embeddings

```
[8]: from sentence_transformers import SentenceTransformer
```

```
# We load our model
```

```
embedding_model = SentenceTransformer('all-mpnet-base-v2')
```

```
.gitattributes: 0%|          | 0.00/1.18k [00:00<?, ?B/s]
1_Pooling/config.json: 0%|          | 0.00/190 [00:00<?, ?B/s]
README.md: 0%|          | 0.00/10.6k [00:00<?, ?B/s]
config.json: 0%|          | 0.00/571 [00:00<?, ?B/s]
config_sentence_transformers.json: 0%|          | 0.00/116 [00:00<?, ?B/s]
data_config.json: 0%|          | 0.00/39.3k [00:00<?, ?B/s]
pytorch_model.bin: 0%|          | 0.00/438M [00:00<?, ?B/s]
sentence_bert_config.json: 0%|          | 0.00/53.0 [00:00<?, ?B/s]
special_tokens_map.json: 0%|          | 0.00/239 [00:00<?, ?B/s]
tokenizer.json: 0%|          | 0.00/466k [00:00<?, ?B/s]
tokenizer_config.json: 0%|          | 0.00/363 [00:00<?, ?B/s]
train_script.py: 0%|          | 0.00/13.1k [00:00<?, ?B/s]
vocab.txt: 0%|          | 0.00/232k [00:00<?, ?B/s]
modules.json: 0%|          | 0.00/349 [00:00<?, ?B/s]
```

```
[9]: import joblib
```

```
[10]: #embeddings = embedding_model.encode(medium_titles)
#joblib.dump(embeddings, model_folder/'medium_nlp_abstract_embeddings')
embeddings_medium = joblib.load(model_folder/'medium_nlp_abstract_embeddings')
```

```
[11]: embeddings_medium.shape
```

```
[11]: (28383, 768)
```

## 0.3 UMAP

```
[12]: from umap import UMAP
      from hdbscan import HDBSCAN
```

```
[13]: umap_model = UMAP(n_neighbors=15, n_components=5, min_dist=0.0,
      ↪metric='cosine', random_state=21)
```

```
##HDBSCAN
```

```
[14]: hdbscan_model = HDBSCAN(min_cluster_size=30, metric='euclidean',
    ↪cluster_selection_method='eom', prediction_data=True)
```

## 0.4 CountVectorizer

```
[15]: from sklearn.feature_extraction.text import CountVectorizer
vectorizer_model = CountVectorizer(stop_words="english", min_df=2,
    ↪ngram_range=(1, 2))
```

```
[67]: # Creating the BERTopic model with embedding model

# topic_model = BERTopic(embedding_model=embedding_model,
    ↪umap_model=umap_model, hdbscan_model=hdbscan_model,
    ↪vectorizer_model=vectorizer_model)
# topics, probs = topic_model.fit_transform(documents=abstracts)
from bertopic import BERTopic
# Creating the BERTopic model without embedding model
topic_model = BERTopic(embedding_model=embedding_model, umap_model=umap_model,
    hdbscan_model=hdbscan_model,
    ↪vectorizer_model=vectorizer_model,
    calculate_probabilities=True)
topics, probs = topic_model.fit_transform(embeddings=embeddings_medium,
    ↪documents=medium_titles)
```

### 0.4.1 Interpreting Topics

#### Overall Summary of Topics

```
[68]: len(topics)
```

```
[68]: 28383
```

```
[94]: topic_model.get_topic_info().head(20)
```

```
[94]:
```

	Topic	Count	Name \
0	-1	10064	-1_job_design_earn_business
1	0	712	0_designs_design_designing_designers
2	1	599	1_newsletters_emails_newsletter_marketing
3	2	447	2_startups_startup_entrepreneurs_entrepreneurship
4	3	396	3_sales_b2b_marketing_b2c
5	4	375	4_data_analyst_analytics_analytical
6	5	367	5_branding_brands_rebranding_rebrands
7	6	362	6_song_jessica_ryan_chris
8	7	359	7_linkedin_jobs_recruiters_company
9	8	300	8_adwords_ads_adsense_advertising
10	9	269	9_marketing_advertising_social_marketer
11	10	262	10_women_men_dating_patriarchy

12	11	258	11_midjourney_guide_style_illustrations
13	12	257	12_seo_marketing_ranking_google
14	13	249	13_products_management_product_managers
15	14	246	14_segmentation_thresholding_opencv_convolutional
16	15	245	15_tensorflow_detector_segmentations_pytorch
17	16	244	16_marketing_advertising_marketer_seo
18	17	240	17_copywriting_copywriters_copywriter_copy
19	18	226	18_diffusion_stablediffusion_denoising_diffusers

#### Representation \

0	[job, design, earn, business, build, model, ai...
1	[designs, design, designing, designers, ux, tr...
2	[newsletters, emails, newsletter, marketing, m...
3	[startups, startup, entrepreneurs, entrepreneu...
4	[sales, b2b, marketing, b2c, business, ecommer...
5	[data, analyst, analytics, analytical, study, ...
6	[branding, brands, rebranding, rebrands, brand...
7	[song, jessica, ryan, chris, music, who, wildf...
8	[linkedin, jobs, recruiters, company, recruite...
9	[adwords, ads, adsense, advertising, ad, campa...
10	[marketing, advertising, social, marketer, bus...
11	[women, men, dating, patriarchy, masculinity, ...
12	[midjourney, guide, style, illustrations, imag...
13	[seo, marketing, ranking, google, wordpress, g...
14	[products, management, product, managers, mana...
15	[segmentation, thresholding, opencv, convoluti...
16	[tensorflow, detector, segmentations, pytorch,...
17	[marketing, advertising, marketer, seo, digita...
18	[copywriting, copywriters, copywriter, copy, w...
19	[diffusion, stablediffusion, denoising, diffus...

#### Representative\_Docs

0	[Make Money Online Listening to Music - Evan C...
1	[Make UX simple again - UX Collective, What th...
2	[The Best 3 Email Scraping Tools That Help Ema...
3	[Is Your Startup a Good Fit for Venture Capita...
4	[B2B sales needs a major overhaul - Sandhya Do...
5	[Becoming a "real" data analyst - Towards Data...
6	[The Brand Archetype Strategy - Umair Asghar -...
7	[If You Try To Do Everything, You Won't Do Any...
8	[5 Tools for LinkedIn Marketing - Sangita Cham...
9	[Digital Marketing (Google Ads Definition - Co...
10	[Developing a Social Media Marketing Strategy ...
11	[Meet the Female Dating Strategy: Women Teachi...
12	[The Ultimate Guide to MidJourney V5.1 (with i...
13	[SEO Guide: To Show You Path for SEO - Chandni...
14	[The Five Whys of a Product Manager - Kubo - M...

```

15 [Introduction to Image Processing with Python,...
16 [DEtection TRansformer (DETR) vs. YOLO for obj...
17 [Digital Marketing - Moose - Medium, What is D...
18 [on copywriting, on copywriting, What About Co...
19 [Stable Diffusion as an API, Stable Diffusion,...

```

```
[70]: dd = topic_model.get_topic_info()
```

```
[71]: dd.shape
```

```
[71]: (169, 5)
```

```
[72]: dd[dd['Topic'] == 126 ]
```

```
[72]:
```

	Topic	Count	Name \
127	126	48	126_father_car_daughter_left

  

	Representation \
127	[father, car, daughter, left, mother, child, m...

  

	Representative_Docs
127	[The Grace Period for Being a Recent Widow Has...

```
[73]: topic_model.get_topic(0)
```

```
[73]: [('ux', 0.06250497792800819),
      ('design', 0.05944338737515278),
      ('collective', 0.04268497037656447),
      ('ux collective', 0.04232014009392014),
      ('ui', 0.028005796164668837),
      ('ux design', 0.027299277203722248),
      ('designer', 0.025242210069824682),
      ('designers', 0.023231444709548417),
      ('ux planet', 0.01865767003122706),
      ('planet', 0.01830805794747912)]
```

```
[74]: topic_model.get_topic(124)
```

```
[74]: [('android', 0.350888540130032),
      ('android apps', 0.07982195451190098),
      ('android interview', 0.053005984616721624),
      ('activity', 0.050965894186584926),
      ('architecture', 0.04843630404000443),
      ('architect', 0.03982334840092273),
      ('interview questions', 0.03974533546622055),
      ('context', 0.037912810350334655),
      ('testing', 0.037887147521482346),
```

```
('app architecture', 0.037092110716688276)]
```

```
[75]: topic_model.visualize_distribution(probs[0])
```

```
[76]: medium_titles[1]
```

```
[76]: 'Generative search experience: the next step in search engine UX?'
```

**Approximate Topic Distribution** For more details, refer to [https://maartengr.github.io/BERTopic/getting\\_started/distribution/distribution.html](https://maartengr.github.io/BERTopic/getting_started/distribution/distribution.html)

```
[77]: topic_distr, _ = topic_model.approximate_distribution(medium_titles)
```

```
[78]: topic_model.visualize_distribution(topic_distr[1])
```

### Interactive Visualizations

```
[79]: # Calculate the topic distributions on a token-level
topic_distr, topic_token_distr = topic_model.
    ↪approximate_distribution(medium_titles, calculate_tokens=True)

# Visualize the token-level distributions
df = topic_model.visualize_approximate_distribution(medium_titles[1],
    ↪topic_token_distr[1])
df
```

```
[79]: <pandas.io.formats.style.Styler at 0x7933e4ddc1f0>
```

```
[80]: topics[1]
```

```
[80]: -1
```

The above document should have been clustered in the generative AI topic rather than search engine. It's appalling that the word generative hasn't been given weightage to model this document into appropriate topic of GEN AI

```
[81]: topic_model.get_document_info(medium_titles)
```

```
[81]:
```

	Document	Topic \
0	5 ChatGPT Prompts I Use to Write High-Ranking ...	-1
1	Generative search experience: the next step in...	-1
2	What Google's Generative AI Announcement Means...	-1
3	Google's Generative AI Updates Will Shatter Or...	-1
4	Will Generative AI Kill SEO? Here's What an SE...	-1
...	...	...
28378	A Successful Freelance Copywriter Addresses th...	61
28379	6 Engaging YouTube Thumbnail Design Ideas That...	-1



28380	What's New in Next.js 14?	49
28381	ChatGPT Has Changed My Approach to Learning Ne...	25
28382	Using ChatGPT For Web Scraping: A Practical Guide	116

	Name \
0	-1_medium_chatgpt_data_life
1	-1_medium_chatgpt_data_life
2	-1_medium_chatgpt_data_life
3	-1_medium_chatgpt_data_life
4	-1_medium_chatgpt_data_life
...	...
28378	61_ai_ai writing_generators_ai write
28379	-1_medium_chatgpt_data_life
28380	49_js_js 13_13_nextjs
28381	25_chatgpt_medium chatgpt_using chatgpt_chat
28382	116_scraping_web scraping_web_scrape

	Representation \
0	[medium, chatgpt, data, life, using, ai, make,...
1	[medium, chatgpt, data, life, using, ai, make,...
2	[medium, chatgpt, data, life, using, ai, make,...
3	[medium, chatgpt, data, life, using, ai, make,...
4	[medium, chatgpt, data, life, using, ai, make,...
...	...
28378	[ai, ai writing, generators, ai write, writing...
28379	[medium, chatgpt, data, life, using, ai, make,...
28380	[js, js 13, 13, nextjs, app, dynamic, material...
28381	[chatgpt, medium chatgpt, using chatgpt, chat,...
28382	[scraping, web scraping, web, scrape, medium w...

	Representative_Docs \
0	[Make Money Online Listening to Music - Evan C...
1	[Make Money Online Listening to Music - Evan C...
2	[Make Money Online Listening to Music - Evan C...
3	[Make Money Online Listening to Music - Evan C...
4	[Make Money Online Listening to Music - Evan C...
...	...
28378	[7 Best Free AI Tools for Content Creation and...
28379	[Make Money Online Listening to Music - Evan C...
28380	[Understanding the SSR, SSG, and ISR in Next.j...
28381	[ChatGPT - twdsmeetup - Medium, What's ...
28382	[Web Scraping and Web Automation with Python 2...

	Top_n_words	Probability \
0	medium - chatgpt - data - life - using - ai - ...	0.517531
1	medium - chatgpt - data - life - using - ai - ...	0.445157
2	medium - chatgpt - data - life - using - ai - ...	0.392761

```

3      medium - chatgpt - data - life - using - ai - ...    0.478484
4      medium - chatgpt - data - life - using - ai - ...    0.392761
...
28378  ai - ai writing - generators - ai write - writ...    0.698075
28379  medium - chatgpt - data - life - using - ai - ...    0.956274
28380  js - js 13 - 13 - nextjs - app - dynamic - mat...    0.456850
28381  chatgpt - medium chatgpt - using chatgpt - cha...    0.245311
28382  scraping - web scraping - web - scrape - mediu...    0.275113

```

```

      Representative_document
0                                False
1                                False
2                                False
3                                False
4                                False
...
28378                           False
28379                           False
28380                           False
28381                           False
28382                           False

```

```
[28383 rows x 8 columns]
```

```
[82]: topic_model.visualize_topics()
```

```
[83]: topic_model.visualize_barchart(top_n_topics=20)
```

```
[84]: topic_model.visualize_barchart(topics=list(range(50, 58, 1)))
```

### 0.4.2 Hierarchical Topics

In BERTopic for creating a hierarchical structure of topics, the default choice is Scipy's ward linkage function. Nonetheless, depending on your specific requirements, you might prefer other linkage functions like single, complete, average, centroid, or median. BERTopic offers the flexibility to customize the linkage function and even select the desired distance function to suit your particular use case.

```

[85]: from scipy.cluster import hierarchy as sch

      # Hierarchical topics
      # linkage_function = lambda x: sch.linkage(x, 'single', optimal_ordering=True)
      # hierarchical_topics = topic_model.hierarchical_topics(docs,
      # ↪linkage_function=linkage_function)

      hierarchical_topics = topic_model.hierarchical_topics(medium_titles)
      topic_model.visualize_hierarchy(hierarchical_topics=hierarchical_topics)

```

100%| | 167/167 [00:01<00:00, 127.61it/s]

Following are the major topics into which the documents are divided:

1. UI/UX
2. ChatGPT
3. Generative AI
4. Google Ads
5. social Media following
6. Email Marketing
7. Big Data

These are just some few examples of clusters which are divided further into subclusters

```
[86]: topic_model.topic_representations_
```

```
[86]: {-1: [('medium', 0.005275896206273433),
        ('chatgpt', 0.00518205698583761),
        ('data', 0.004786524368214274),
        ('life', 0.004722417685825071),
        ('using', 0.004604612433825728),
        ('ai', 0.004364726776236341),
        ('make', 0.004253299534447538),
        ('google', 0.0040877865016399),
        ('science', 0.0038601033299912177),
        ('data science', 0.00364780686384404)],
 0: [('ux', 0.06250497792800819),
      ('design', 0.05944338737515278),
      ('collective', 0.04268497037656447),
      ('ux collective', 0.04232014009392014),
      ('ui', 0.028005796164668837),
      ('ux design', 0.027299277203722248),
      ('designer', 0.025242210069824682),
      ('designers', 0.023231444709548417),
      ('ux planet', 0.01865767003122706),
      ('planet', 0.01830805794747912)],
 1: [('email', 0.10096506389619882),
      ('email marketing', 0.06558259092129835),
      ('newsletter', 0.057532295180489156),
      ('emails', 0.03268372462660052),
      ('substack', 0.03162205228756058),
      ('email list', 0.030878460859256326),
      ('list', 0.028821951404442398),
      ('subscribers', 0.026802154709590704),
      ('marketing', 0.023475923997711808),
      ('newsletters', 0.021717214134756677)],
 2: [('startup', 0.07255206328241741),
      ('founders', 0.037222653584189684),
      ('venture', 0.0363147839845753),
```

```

('entrepreneur', 0.03528028087852098),
('startups', 0.03421447131844151),
('founder', 0.028633134277186497),
('investors', 0.028277800439257136),
('entrepreneurs', 0.026795447570054573),
('entrepreneur handbook', 0.025741600371241492),
('vc', 0.023945550590851013)],
3: [('b2b', 0.06245900344237648),
('sales', 0.05451236743660781),
('marketing', 0.03425345547509847),
('b2b sales', 0.02665099922440018),
('funnel', 0.026520688265398506),
('b2b marketing', 0.022806192672396166),
('customers', 0.018004755887214934),
('sales funnel', 0.016815548623647557),
('marketing strategy', 0.015789310013528762),
('b2c', 0.01479872614573983)],
4: [('data', 0.06451462529891962),
('data engineering', 0.03585864754166752),
('data analyst', 0.03261652197471509),
('data science', 0.03197610186455185),
('analyst', 0.031522291427078034),
('data engineer', 0.0314010401255264),
('science', 0.031114505992730017),
('engineering', 0.030771582106318047),
('engineer', 0.027758255453537384),
('science data', 0.020697696521423403)],
5: [('brand', 0.12914703122480167),
('branding', 0.06491816029856762),
('identity', 0.03922589423985105),
('brand identity', 0.038664229637885926),
('logo', 0.036706112628214466),
('brand strategy', 0.03335104059990857),
('brands', 0.03284846426480009),
('brand building', 0.019723437363290432),
('medium brand', 0.01826044678098825),
('strategy', 0.01655276309593691)],
6: [('medium', 0.028948625372405248),
('jessica', 0.017401199386835733),
('holiday medium', 0.014751824872191394),
('ryan holiday', 0.014465215767477524),
('holiday', 0.01334093596436887),
('owl', 0.012916200797696554),
('ryan', 0.012673735733755768),
('jessica wildfire', 0.011636496331383264),
('wildfire', 0.011636496331383264),
('wildfire medium', 0.011636496331383264)],

```

```

7: [('linkedin', 0.18487280480325807),
    ('profile', 0.02593926971056211),
    ('leads', 0.02218253155509211),
    ('recruiters', 0.019659630279732707),
    ('using linkedin', 0.01750992566558616),
    ('job', 0.013367546205797108),
    ('tips', 0.012963859052820003),
    ('posts', 0.012693190640822308),
    ('engagement', 0.011508325074064664),
    ('linkedin lead', 0.011494361449713743)],
8: [('google ads', 0.10731395438736328),
    ('ads', 0.10491145442795342),
    ('google', 0.07415405024489266),
    ('adwords', 0.0370077797177708),
    ('google adwords', 0.03515963969950963),
    ('ad', 0.02723227508657553),
    ('campaign', 0.024987036198124864),
    ('medium google', 0.019070798634614496),
    ('ads campaign', 0.01628482847737962),
    ('search ads', 0.01599048769167791)],
9: [('social media', 0.11895742769054993),
    ('social', 0.11344249649756576),
    ('media', 0.1116495496995904),
    ('media marketing', 0.08102564149838203),
    ('medium social', 0.028542840767762534),
    ('marketing', 0.02242308948101716),
    ('management', 0.019684734373327083),
    ('media strategy', 0.017976052260299233),
    ('media manager', 0.01767189756943169),
    ('media optimization', 0.016642674376952316)],
10: [('men', 0.10645425626876043),
    ('women', 0.09737586807565352),
    ('relationship', 0.04449002628245676),
    ('dating', 0.044151140719372116),
    ('woman', 0.04168429595739842),
    ('trans', 0.03627444447429048),
    ('gender', 0.0301443239434344),
    ('relationships', 0.025541345937881817),
    ('man', 0.025455380915653535),
    ('love', 0.024355779176851878)],
11: [('midjourney', 0.15240738415512853),
    ('midjourney prompts', 0.05306043502914712),
    ('prompts', 0.0511255530513798),
    ('midjourney ai', 0.04212312072625517),
    ('images', 0.03750176878065567),
    ('v5', 0.03403939225474404),
    ('midjourney v5', 0.029744196805259703),

```

```

('prompts create', 0.02396184311615709),
('using midjourney', 0.02355362901113372),
('images midjourney', 0.022987217587099866)],
12: [('seo', 0.13307274526133425),
('ranking', 0.032706323119819074),
('medium seo', 0.024294243665684254),
('rank', 0.0234908960720873),
('website', 0.022078688467174087),
('seo tips', 0.01988630260467584),
('rankings', 0.01951856742066837),
('nolan medium', 0.019103439464679975),
('nick nolan', 0.019103439464679975),
('nolan', 0.01894559140635326)],
13: [('product', 0.14051984090859224),
('product manager', 0.07501050685087668),
('product management', 0.06827418742742679),
('manager', 0.06602560260600329),
('management', 0.04936104966326129),
('product managers', 0.045640941655307375),
('managers', 0.03989920577174291),
('coalition', 0.03435729532672863),
('product coalition', 0.03435729532672863),
('product development', 0.026942467470898877)],
14: [('image', 0.08277555753258073),
('image processing', 0.060203185898455386),
('opencv', 0.05453110655848541),
('computer vision', 0.048091818804974006),
('pytorch', 0.04527792728549669),
('computer', 0.04510538100137988),
('processing', 0.041642299081386606),
('vision', 0.03899502938060364),
('python', 0.038401289400317486),
('python image', 0.026081650723824845)],
15: [('object', 0.12521689379160397),
('object detection', 0.12139535384251286),
('detection', 0.11300059120500171),
('yolov8', 0.08154440543089063),
('yolo', 0.07712860716908747),
('yolov7', 0.040263447909722964),
('segmentation', 0.032714877072262435),
('custom', 0.030708009771913694),
('model', 0.027042982755837715),
('training', 0.026143326196752623)],
16: [('digital marketing', 0.147592349264416),
('digital', 0.12850383591217449),
('marketing', 0.05846930199370278),
('medium digital', 0.056948594685003066),

```

```

('best digital', 0.029735261212091902),
('marketing company', 0.026795722333329844),
('marketing agency', 0.022551139580826614),
('agency', 0.021857923329000533),
('company', 0.02054125472742188),
('marketing strategies', 0.015897495285960617)],
17: [('copywriting', 0.17491818267317777),
('copywriter', 0.07839731957585266),
('copy', 0.07263865980583052),
('copywriter handbook', 0.048063398165611476),
('handbook medium', 0.04764594012497636),
('copywriting tips', 0.041209967498541486),
('handbook', 0.03636977023915266),
('benjamin watkins', 0.031547054828555646),
('watkins medium', 0.03134941353218513),
('watkins', 0.03134941353218513)],
18: [('diffusion', 0.1828017697918788),
('stable diffusion', 0.16991276658747703),
('stable', 0.1667095891658714),
('controlnet', 0.03904397326481505),
('sdxl', 0.03453601174102895),
('image', 0.03301714121371842),
('image generation', 0.032645463677631124),
('dreambooth', 0.023726776448892237),
('lora', 0.020104680595895326),
('models', 0.019633885569800252)],
19: [('ai art', 0.07132616906884702),
('art', 0.059638703791676996),
('ai', 0.05794213941206175),
('ai image', 0.03291351592903616),
('mlearning', 0.032298890025931046),
('mlearning ai', 0.032298890025931046),
('images', 0.0315319770006113),
('image', 0.03146018570421566),
('generator', 0.028802374832310327),
('ai generated', 0.028577586793999948)],
20: [('apple', 0.09175984727018575),
('apps', 0.06875428876934314),
('iphone', 0.05804490088254749),
('mac', 0.05692121295130406),
('macos', 0.05199988279341018),
('productivity', 0.05039098980408355),
('pro', 0.0393063588799523),
('m2', 0.034728186602936646),
('macbook', 0.0335896840877216),
('ios', 0.03185176525625294)],
21: [('instagram', 0.13259904204990045),

```

```

('followers', 0.046277001222050994),
('grow instagram', 0.03539066462191029),
('instagram followers', 0.03486229043614734),
('instagram algorithm', 0.034450356950004576),
('instagram growth', 0.028778409480338128),
('medium instagram', 0.02801980246963183),
('grow', 0.027825250956378656),
('algorithm', 0.0276186220777164),
('followers instagram', 0.02493538522293245)],
22: [('sentiment', 0.17456081860408887),
('sentiment analysis', 0.16230054666760155),
('analysis', 0.11356880648713318),
('text', 0.03517831502289532),
('reviews', 0.03433642995546031),
('text mining', 0.03243966237272199),
('analysis using', 0.029028527521168995),
('nlp', 0.02896994268122454),
('text data', 0.02652292025637781),
('mining', 0.02533885644550572)],
23: [('landing', 0.1800484730173215),
('landing page', 0.16938412943476297),
('page', 0.14956019149847424),
('landing pages', 0.053630842712471496),
('pages', 0.050714978369593676),
('converting', 0.02948126698693277),
('high converting', 0.029441317827487396),
('conversion', 0.02727130482079701),
('designing landing', 0.026255367648938797),
('case study', 0.02602074456099804)],
24: [('language', 0.1605653996295301),
('language models', 0.14245508428134548),
('large language', 0.1419568179036281),
('large', 0.13287054252485875),
('models', 0.09742886603429152),
('language model', 0.07361253231576276),
('model', 0.03980946700097938),
('llms', 0.025007415016616238),
('open', 0.01849252472433433),
('models large', 0.01766617876507047)],
25: [('chatgpt', 0.09819670590844437),
('medium chatgpt', 0.044099641109814354),
('using chatgpt', 0.023087357241932907),
('chat', 0.02086031916090708),
('chatgpt plugins', 0.020566615275311333),
('use chatgpt', 0.018999399955047687),
('chat gpt', 0.018167230024145914),
('plugins', 0.016626826834008058),

```



```

('chatgpt wrong', 0.01596345536303953),
('people chatgpt', 0.01596345536303953)],
26: [('twitter', 0.17420615289188182),
('followers', 0.054542388366695956),
('medium twitter', 0.02730446205363221),
('grow', 0.025580123411400466),
('twitter marketing', 0.02183517740443083),
('twitter account', 0.019219363397112985),
('100 followers', 0.018670800518878126),
('tweet', 0.017065288783520132),
('days', 0.01659055983812563),
('twitter api', 0.014936640415102502)],
27: [('affiliate', 0.20144834791794083),
('affiliate marketing', 0.15587236003946842),
('marketing', 0.04203223238011116),
('medium affiliate', 0.03304446775081352),
('programs', 0.029875305792347425),
('day affiliate', 0.026089363727126405),
('affiliate program', 0.018545308633896904),
('ticket', 0.01701626912156295),
('marketer', 0.016793509613280345),
('promote', 0.016682798069023324)],
28: [('python', 0.10479718076709357),
('libraries', 0.03732396177992128),
('python libraries', 0.03245887580865462),
('coding', 0.03049269455952533),
('code', 0.030435868338538508),
('python tricks', 0.030056445091339222),
('python code', 0.02914525545537559),
('level', 0.02726091708902439),
('level coding', 0.025037908440036876),
('decorators', 0.02367508310693118)],
29: [('llm', 0.1427456197605585),
('llms', 0.10241576203233611),
('llama', 0.06261666353277762),
('fine', 0.045625814850631836),
('tuning', 0.03893427333328379),
('fine tuning', 0.0386502820227607),
('langchain', 0.029313424349840963),
('open source', 0.02525918171074991),
('source', 0.024314933767649705),
('lora', 0.022988209056628572)],
30: [('storytelling', 0.17661254451734776),
('narrative', 0.03974402497424491),
('hollywood', 0.03445859961346015),
('art storytelling', 0.03109391621882178),
('screenwriting', 0.023325986707231117),

```

```

('stories', 0.023262280640695326),
('art', 0.020681747512276066),
('speaking', 0.02058060375804722),
('storytellers', 0.01985802490201164),
('story', 0.019734973408500913)],
31: [('leadership', 0.1088199163297895),
('meetings', 0.06617910429287459),
('leader', 0.0548067429424913),
('team', 0.04836896692506439),
('leaders', 0.04198743008733073),
('manager', 0.03998306294605228),
('boss', 0.03574016727374525),
('communication', 0.03304437163240539),
('employee', 0.027532448510407613),
('layoffs', 0.024427595900217033)],
32: [('spark', 0.21276292859578827),
('apache spark', 0.1289039631163645),
('apache', 0.10844396668449229),
('medium spark', 0.031776123619408174),
('spark spark', 0.02996584540756729),
('spark data', 0.029845838117725208),
('optimization', 0.029418171472257983),
('partitioning', 0.026220114731621377),
('medium apache', 0.025668397611070027),
('alsathamhussain medium', 0.023603718877520684)],
33: [('crypto', 0.14094825078707163),
('bitcoin', 0.09720666644336351),
('blockchain', 0.0519124788513265),
('coinmonks', 0.047869538582560864),
('coinmonks medium', 0.04486441076966636),
('ethereum', 0.037406889370551914),
('bull', 0.03333351617475609),
('cryptocurrency', 0.0311882441438707),
('cryptozoa', 0.026049679765859902),
('market', 0.021914996375572423)],
34: [('books', 0.20966739609991752),
('read', 0.1448874872024223),
('10 books', 0.07794889392620208),
('read books', 0.06389712918106837),
('books read', 0.05537751195692591),
('book', 0.045358955325107045),
('reading', 0.03970208584911871),
('favorite books', 0.029956406291849168),
('books superpower', 0.029267426537665696),
('updated 2023', 0.028094924890317344)],
35: [('habits', 0.06585320610860966),
('fitness', 0.05928822405834935),

```

```

('life', 0.04159168411237871),
('exercise', 0.035376386288025405),
('fitness health', 0.03394545585254787),
('brain', 0.0322572027218665),
('weight', 0.031996575859152575),
('weight loss', 0.030558180439482874),
('health', 0.030518026658581898),
('fat', 0.029096105016469604)],
36: [('lead', 0.1248524084208588),
('lead generation', 0.12461277647151382),
('leads', 0.10355014730214301),
('generation', 0.08919751896085404),
('real estate', 0.03792121307414664),
('estate', 0.037515821899963706),
('generation strategies', 0.03287608223448229),
('generate', 0.024760508401619724),
('generate leads', 0.02140647242225659),
('leads month', 0.019178208443371845)],
37: [('confirmed', 0.0816270982957725),
('claim', 0.07055733769558063),
('tokens', 0.056632834674754114),
('000', 0.0476521119106976),
('token', 0.04633595564298063),
('farm', 0.04194586166228037),
('network', 0.03648531563646965),
('free', 0.03628618410842199),
('early access', 0.031253347208651236),
('horizonindex', 0.031253347208651236)],
38: [('chatbot', 0.12973183081661724),
('bot', 0.07834134624094813),
('chatbot using', 0.03562000831932953),
('telegram', 0.03463360352952588),
('gpt', 0.03429271739057383),
('build custom', 0.033150661186626205),
('ai chatbot', 0.03013696471511473),
('build', 0.029857613611162105),
('custom', 0.029842305241109743),
('chatgpt', 0.02805067446163206)],
39: [('visualization', 0.09629714634691346),
('data visualization', 0.08399684986303785),
('dashboard', 0.05539548439640988),
('bi', 0.05070605217033073),
('power bi', 0.04830146673109882),
('charts', 0.04724110536905025),
('plots', 0.04181411167950347),
('dashboards', 0.03912304711217499),
('tableau', 0.03530749361755214),

```

```

('split', 0.03477604187748888)],
40: [('content', 0.1329841797153902),
('content marketing', 0.0946938047173342),
('audience', 0.05018137617945739),
('content strategy', 0.03968409339142127),
('create content', 0.03511764386836239),
('strategy', 0.03151356646242852),
('marketing strategy', 0.02723205654284241),
('best content', 0.024470555582116537),
('marketing', 0.023419671235729037),
('make content', 0.01957644446569323)],
41: [('writing', 0.10481274990180436),
('write', 0.05894428223476919),
('writer', 0.0562842010042501),
('better writer', 0.034084008733578676),
('writers', 0.0318498887354088),
('writing cooperative', 0.028418688088155376),
('cooperative', 0.02821691510562049),
('improve writing', 0.025395576988972605),
('new writers', 0.024804257004003297),
('struggling', 0.024804257004003297)],
42: [('machine learning', 0.10160796390839091),
('machine', 0.09461604597557297),
('learning', 0.07274159512215282),
('logistic regression', 0.06987114649015014),
('logistic', 0.06896342135919858),
('regression', 0.062157688781315786),
('learning algorithms', 0.035828143205861855),
('algorithms', 0.03193564335394673),
('xgboost', 0.03191971474051589),
('overfitting', 0.03151120216345065)],
43: [('customer', 0.1560626995973214),
('segmentation', 0.07899609570320303),
('churn', 0.06627588450865105),
('retention', 0.052636572350670145),
('rfm', 0.04703640983460925),
('lifetime', 0.043907130263241126),
('analysis', 0.04177673224897951),
('customer retention', 0.03359733428927046),
('value', 0.032140510907644686),
('crm', 0.03150960679915623)],
44: [(' ', 0.13767752807489309),
('kullanımı', 0.09677493028974643),
(' ', 0.08511449714801615),
('webinar', 0.07779545778052548),
('nasıl', 0.07301269029664118),
('smo', 0.06258069457949686),

```

```

('rs', 0.05841015223731294),
('linkedin', 0.05738168366997767),
(' ', 0.05674299809867744),
('smm', 0.050130934008910936)],
45: [('youtube', 0.21877805532394623),
('channel', 0.08819455554936118),
('faceless youtube', 0.07350251731721008),
('faceless', 0.0733820653334157),
('videos', 0.06896391391275036),
('youtube videos', 0.05779252543084818),
('youtube shorts', 0.04014725158293736),
('shorts', 0.037594624626467205),
('youtube channels', 0.03260862944714572),
('watching', 0.03238425112626331)],
46: [('rich', 0.08040859088832152),
('financial', 0.06061589396931574),
('wealthy', 0.045158678493124126),
('dad', 0.04303011446768805),
('wealth', 0.03822674271334071),
('money', 0.03712927404115466),
('financial freedom', 0.03386900886984309),
('millionaire', 0.03340766065788213),
('afford', 0.032272585850766033),
('poor', 0.03181158809816976)],
47: [('figma', 0.2554805731023967),
('plugins', 0.09485468423335744),
('layout', 0.05410577483910021),
('responsive', 0.05224953954231553),
('auto layout', 0.048887855982297425),
('ux collective', 0.043621669125863634),
('collective', 0.043293756019640056),
('auto', 0.03780365515426071),
('ux', 0.03594499027906407),
('tips tricks', 0.03324184854750777)],
48: [('interview', 0.15684803338838108),
('internship', 0.10679758065978058),
('interview experience', 0.07361925989746433),
('interviews', 0.06295838845563617),
('questions', 0.058103211714533136),
('experience', 0.05151845289803821),
('interview questions', 0.0490442441413363),
('design interview', 0.04201280906554164),
('preparation', 0.035360097310653794),
('intern', 0.03283242657281965)],
49: [('js', 0.20050858867920207),
('js 13', 0.10667354801570444),
('13', 0.09376686876411802),

```

```

('nextjs', 0.08760290333811283),
('app', 0.05410482256719433),
('dynamic', 0.03836191322782176),
('material ui', 0.035633828157531224),
('routes', 0.035633828157531224),
('material', 0.0339289051671005),
('dynamic web', 0.029694856797942686)],
50: [('nlp', 0.1395628875672962),
('bert', 0.10227354120965225),
('language processing', 0.05994822193556431),
('natural language', 0.059659565705630475),
('natural', 0.05748498420819075),
('text', 0.054687696171603586),
('nlp data', 0.0509824833974737),
('processing', 0.046444524209162126),
('using nlp', 0.04363411309580095),
('spacy', 0.041577409788336524)],
51: [('nft', 0.16305560707148997),
('games', 0.1361932228775132),
('nfts', 0.0749280221889682),
('soon', 0.05198496976886919),
('sale', 0.04949097399025721),
('arena', 0.03907451964878985),
('earn games', 0.03907451964878985),
('mint', 0.03808081089245796),
('play earn', 0.03720497460314539),
('blog', 0.03660687350951341)],
52: [('trading', 0.13146600113837703),
('stock', 0.07305421894601156),
('python', 0.04656072690709285),
('market', 0.04123040073575223),
('trading bot', 0.039527511785131),
('bot', 0.0390408829510888),
('pricing', 0.03803649096630051),
('bot python', 0.03458657281198963),
('options', 0.03451884672525309),
('dynamic pricing', 0.031135323012395574)],
53: [('hashtags', 0.22957608279015956),
('hashtag', 0.12518056629550398),
('instagram hashtags', 0.04779408275942523),
('instagram', 0.03997186762486746),
('trending', 0.035172663232262105),
('captions', 0.030574497206925386),
('using hashtags', 0.026708088780469985),
('tags', 0.02179134371708296),
('twitter', 0.021366174352790812),
('hashtags boost', 0.020832880973158618)],

```

```

54: [('pyspark', 0.27800869968538805),
    ('subham khandelwal', 0.03167691641431022),
    ('subham', 0.03167691641431022),
    ('analysis pyspark', 0.03167691641431022),
    ('khandelwal', 0.0279486597151151),
    ('khandelwal medium', 0.0279486597151151),
    ('data', 0.026273921030266965),
    ('json', 0.02599728490254983),
    ('functions', 0.025543748540606606),
    ('count', 0.023162401669282663)],
55: [('growth', 0.20410674310949428),
    ('growth hacking', 0.10761319200664297),
    ('hacking', 0.10356127232299477),
    ('growth marketing', 0.10070590205560459),
    ('growth hackers', 0.053537840144779844),
    ('hackers', 0.05306534200122955),
    ('medium growth', 0.04083151981308753),
    ('growth hacker', 0.03190926543446409),
    ('hacker', 0.028153662109113502),
    ('growth strategy', 0.02716519543441446)],
56: [('photography', 0.12832153029541316),
    ('paintings', 0.07042737629331704),
    ('minimalism', 0.05502492184848956),
    ('art', 0.05184882484528565),
    ('street', 0.049421118553756097),
    ('camera', 0.04187461254686675),
    ('artist', 0.04187461254686675),
    ('paris', 0.039658553548581485),
    ('perception', 0.03591140076952022),
    ('photos', 0.03551796680544334)],
57: [('ai', 0.05143899034354422),
    ('ai tools', 0.048942135195465725),
    ('ai marketing', 0.03593660378065256),
    ('tools', 0.03584887000940815),
    ('tools business', 0.027112684186845565),
    ('marketing', 0.019770159255647866),
    ('ai startups', 0.017490370302389473),
    ('powerful ai', 0.017490370302389473),
    ('business', 0.01726683344583844),
    ('startups', 0.01600864385563568)],
58: [('airbnb', 0.11893315351132583),
    ('case study', 0.08762158053264729),
    ('study', 0.08254166925508495),
    ('booking', 0.08156576954112073),
    ('case', 0.07737754796510313),
    ('ux case', 0.07387996314102402),
    ('app', 0.053859727108226285),

```

```

('travel', 0.03998257298891015),
('booking com', 0.03782580204655221),
('ux', 0.03585463881329065)],
59: [('google analytics', 0.19610603384400702),
('analytics', 0.19289881075474302),
('ga4', 0.10065617020952694),
('google', 0.08179243036752548),
('analytics google', 0.05225144491072446),
('updates google', 0.038813054461548875),
('universal analytics', 0.038813054461548875),
('universal', 0.037652062986250366),
('tracking', 0.03547211723198115),
('data analytics', 0.033806663863019316)],
60: [('white', 0.21814537534817954),
('black', 0.21158472278422222),
('privilege', 0.05666012472420794),
('black women', 0.05601240155663438),
('woke', 0.052210244681008475),
('women', 0.051578562453520334),
('men', 0.046665311094874706),
('asian', 0.04612647215307117),
('action', 0.04207535681931562),
('america', 0.03966599338696202)],
61: [('ai', 0.05446408330306587),
('ai writing', 0.05303252498515009),
('generators', 0.03953067490059903),
('ai write', 0.03845915156223184),
('writing', 0.03345950184025991),
('content creation', 0.031380640195469045),
('ai generated', 0.029654304127011227),
('generated', 0.0288201485903062),
('best ai', 0.028594722822680434),
('creation', 0.028070679692125528)],
62: [('en', 0.12606395935341583),
('cómo', 0.1078098735472494),
('para', 0.105867101543909),
('que', 0.0994292734080765),
('la', 0.09112731011266723),
('el', 0.0650244294557237),
('da', 0.06056925826039985),
('los', 0.06056925826039985),
('se', 0.059841135922414156),
('medium cómo', 0.059841135922414156)],
63: [('bigquery', 0.27739695353911653),
('google bigquery', 0.07577391772874316),
('google cloud', 0.04271031647629705),
('sql', 0.041690689834186084),

```



```

('ga4', 0.040162931697942265),
('google', 0.03852871694176906),
('cloud', 0.03817420638333271),
('cloud community', 0.03608896022944368),
('bigquery data', 0.03321220028326818),
('bigquery sql', 0.03321220028326818)],
64: [('writing', 0.06139587300808614),
('money writing', 0.054970676772721536),
('pay', 0.04955118504545717),
('websites pay', 0.04500772440242534),
('articles', 0.03723999130974017),
('money', 0.0355037989103207),
('sites', 0.034282568560969874),
('500', 0.034042863510635145),
('make', 0.029366510755740528),
('article', 0.028981101379338743)],
65: [('hustles', 0.1512985602178788),
('hustle', 0.13306784451815987),
('hustles make', 0.06917513816591127),
('month', 0.05131401884385804),
('club medium', 0.03289030900876995),
('hustle club', 0.03174666703461328),
('time job', 0.028847973058477805),
('club', 0.028326206510120966),
('hustles earn', 0.026869564528840418),
('500', 0.026797949389871192)],
66: [('ppc', 0.2697780414233774),
('pay click', 0.05881729431678613),
('advertising', 0.05506192842921506),
('click', 0.046814153870166304),
('campaigns', 0.0327279232927316),
('pay', 0.03193699978493486),
('click advertising', 0.03171540598467998),
('campaign', 0.027371536054025924),
('seo ppc', 0.027000157795080224),
('parashar medium', 0.023650035500704686)],
67: [('forecasting', 0.21261658540448528),
('time series', 0.20962589141928967),
('series forecasting', 0.18374299444906572),
('series', 0.17667824609221328),
('time', 0.10570179640205338),
('temporal', 0.055697915298119405),
('fusion', 0.05469120109332639),
('deep learning', 0.05288432628942513),
('lstm', 0.048799870721359766),
('forecast', 0.047030863742186825)],
68: [('100 day', 0.08999454334039574),

```

```

('websites', 0.08156728311975657),
('100', 0.06711411285787008),
('day', 0.06573857488860216),
('make 100', 0.053513140131204555),
('earn', 0.04993605524600852),
('websites pay', 0.0386703338592479),
('earn 100', 0.038167951464946205),
('make', 0.03500871787050344),
('pay', 0.03459141128761066)],
69: [('saas', 0.21217354467638666),
('micro saas', 0.08033161034887897),
('micro', 0.06910529205864863),
('saas ideas', 0.06908908681576452),
('build saas', 0.0486599480426999),
('ideas', 0.04850467237484651),
('saas startup', 0.044516467933283364),
('ideas build', 0.04349261531107696),
('saas business', 0.03966024556509099),
('vertical saas', 0.038156972514242875)],
70: [('passive income', 0.21685014569934985),
('passive', 0.2128028544786552),
('income', 0.18324426711876377),
('month', 0.04957557661667126),
('income streams', 0.04563867380993597),
('ideas', 0.04055973587665472),
('streams', 0.040480523849755284),
('income 2023', 0.027870646868412553),
('income month', 0.025959263861391385),
('medium passive', 0.025959263861391385)],
71: [('flutter', 0.3933903577911796),
('bloc', 0.04939093755361021),
('responsive design', 0.03740879658259105),
('mvvm', 0.03654841622779576),
('app', 0.03617505504809621),
('desktop', 0.03577937356080875),
('dart', 0.031862053253066125),
('riverpod', 0.031862053253066125),
('state management', 0.029634562532166127),
('responsive', 0.028843637213404955)],
72: [('openai', 0.19562870692972226),
('openai gpt', 0.10990529872075741),
('openai api', 0.07246086675794991),
('api', 0.07142643691687174),
('gpt api', 0.07055481535928132),
('gpt', 0.06695578873048624),
('guide openai', 0.03722631464804183),
('azure', 0.03507067329465475),

```

```

('use openai', 0.03183143795779466),
('azure openai', 0.028568371302814803)],
73: [('chatgpt', 0.06965960980723065),
('make money', 0.04377308623978466),
('money', 0.03721230014589785),
('make', 0.03457963553330997),
('using chatgpt', 0.03114108651237462),
('chatgpt make', 0.0303509059597577),
('asked chatgpt', 0.02635232263461195),
('chat gpt', 0.025729867638848516),
('asked', 0.024783739611763406),
('way make', 0.021806057838559118)],
74: [('gpt', 0.17462098094566253),
('tuning gpt', 0.04868831371298193),
('medium gpt', 0.04449792713031239),
('fine', 0.04117497005499877),
('tuning', 0.03593474513260734),
('fine tuning', 0.030828201137201985),
('autogpt', 0.03043939256605437),
('gpt human', 0.03043939256605437),
('ai gpt', 0.025266701121964972),
('gpt4', 0.025266701121964972)],
75: [('software', 0.1467508931370409),
('software engineer', 0.12482050714029161),
('engineer', 0.08089589676572975),
('principal', 0.0618246861886036),
('programming', 0.06144633146246012),
('better programming', 0.060207905350334136),
('programmer', 0.05546974937163056),
('software engineering', 0.050981904077020204),
('developer', 0.043525267479077055),
('career', 0.0423779973877224)],
76: [('lin medium', 0.10235991813879067),
('lin', 0.10235991813879067),
('product designer', 0.09414715621543739),
('dean', 0.07796995461929764),
('designer', 0.0560590671519739),
('ching', 0.04756517727198849),
('man wongnai', 0.04756517727198849),
('wongnai', 0.04756517727198849),
('line man', 0.04756517727198849),
('medium', 0.04666948925458804)],
77: [('facebook', 0.20345238035739177),
('facebook ads', 0.1817926583060182),
('ads', 0.11591809319947959),
('ad', 0.06526835664921579),
('days facebook', 0.03843206742388951),

```

```

('ads best', 0.03745469694649608),
('medium facebook', 0.0358232746326625),
('fb', 0.031212247455413396),
('ads vs', 0.029272728490013677),
('ad campaign', 0.023418182792010942)],
78: [('narcissistic', 0.09425655795196107),
('subtle', 0.0829898799407305),
('behaviours', 0.07171272431284376),
('signs', 0.06816677867443535),
('attractive', 0.058910348719975664),
('ex', 0.056988881273942964),
('abuse', 0.05576197483112553),
('respect', 0.05576197483112553),
('attractive people', 0.053804188535958905),
('subtle behaviours', 0.053804188535958905)],
79: [('laravel', 0.3721638435310917),
('php', 0.1283701139158116),
('crud', 0.04654105437247401),
('laravel application', 0.038109335844437334),
('creates', 0.03723284349797921),
('know 2023', 0.033952056128488804),
('solid', 0.0334410065619673),
('medium laravel', 0.03245872091698122),
('hidden features', 0.03245872091698122),
('features need', 0.030189516886701073)],
80: [('generative ai', 0.22859605275122422),
('generative', 0.22024543310992473),
('medium generative', 0.08642786127833281),
('ai', 0.06329418103296704),
('ai generative', 0.0597311989761653),
('ai change', 0.032018919114338656),
('ai landscape', 0.030972530292984385),
('landscape', 0.028827479388732626),
('change 2023', 0.024975442319406297),
('like midjourney', 0.024975442319406297)],
81: [('etsy', 0.28696190961960794),
('print demand', 0.13521903388853143),
('print', 0.1325655338594587),
('demand', 0.10857835631533296),
('selling', 0.05215390635230594),
('tutorial', 0.04098901604776926),
('beginner', 0.039025544273701404),
('designs', 0.036154236507125105),
('success tips', 0.03413384003894167),
('products', 0.03243967520958822)],
82: [('ai', 0.06830474667877726),
('intelligence', 0.05622035847463052),

```

```

('replaced', 0.04643124813234723),
('replaced ai', 0.04102087027353275),
('ai changing', 0.04102087027353275),
('artificial', 0.04059151816974131),
('human', 0.037770432622732955),
('work', 0.03486962796991122),
('artificial intelligence', 0.03426855072508665),
('crisis', 0.0302381225147679)],
83: [('spring', 0.15915393459980304),
('authentication', 0.141578609134832),
('spring boot', 0.1406667740831535),
('boot', 0.13916299653778344),
('sso', 0.08896398213088583),
('keycloak', 0.05674299809867744),
('login', 0.05504567903968124),
('jwt', 0.05289124787249989),
('net core', 0.04502341129511073),
('server', 0.04144016121515929)],
84: [('pandas', 0.18705966437912774),
('cleaning', 0.07970582231317641),
('data', 0.04347172797969037),
('league', 0.04097610537738088),
('mastering data', 0.03993403716822907),
('python pandas', 0.03374401259549634),
('dataframe', 0.03321075929715684),
('dataset', 0.029934572970310314),
('hands examples', 0.027210293515962273),
('pivot tables', 0.027210293515962273)],
85: [('kafka', 0.2983443727111049),
('streaming', 0.10894489904440133),
('apache kafka', 0.08224891287591014),
('apache', 0.06205325295341572),
('consumer', 0.05845988190325454),
('using apache', 0.05797888800029396),
('spark streaming', 0.049554509758756984),
('flink', 0.04739605328834406),
('spark', 0.04722273618843356),
('lag', 0.04361797082656486)],
86: [('freelancer', 0.20961671566966208),
('freelance', 0.12103056578933852),
('freelancing', 0.0844521104682192),
('clients', 0.07402981200620733),
('freelancers', 0.07237683821402403),
('upwork', 0.06675484644252538),
('freelance work', 0.03459672049833926),
('freelance clients', 0.033265164469088365),
('client', 0.03305749570594117),

```

```

('fiverr', 0.02533563314046576)],
87: [('ai tools', 0.15716231156808258),
('tools', 0.10510718987965526),
('free ai', 0.06175047323092788),
('ai', 0.05856179778457477),
('productivity', 0.03632723229465983),
('best ai', 0.033574023669233406),
('free', 0.0261128308715466),
('tools try', 0.025640468890041377),
('tools best', 0.024653620412741526),
('tools start', 0.024653620412741526)],
88: [('aws', 0.18770236146563643),
('terraform', 0.13165699088496993),
('sagemaker', 0.09742183990988737),
('serverless', 0.07504361960218053),
('deploying', 0.05566962280564993),
('aws sagemaker', 0.05308796175894662),
('endpoints', 0.04423996813245551),
('cloud', 0.04131496048711192),
('offerings', 0.038645443244088744),
('cost', 0.031229079269470272)],
89: [('angular', 0.21661207829090195),
('node js', 0.11521089765468395),
('node', 0.10420320878813036),
('js', 0.09947733171984119),
('nodejs', 0.0694759873287602),
('express', 0.06627145885681429),
('express typescript', 0.05500104146197172),
('development nodejs', 0.047030232658363286),
('nodejs express', 0.047030232658363286),
('api development', 0.047030232658363286)],
90: [('medium medium', 0.6082062755425067),
('medium', 0.061240262973474395),
('brandeis', 0.05589456315371919),
('marshall medium', 0.05589456315371919),
('brandeis marshall', 0.05589456315371919),
('50', 0.0544234979357667),
('wise medium', 0.05325025182693693),
('marshall', 0.05325025182693693),
('center', 0.05120076005841251),
('medium 50', 0.04811387857215046)],
91: [('react', 0.3254982202021407),
('react native', 0.04777739723957889),
('senior', 0.043825443136800375),
('native', 0.043609324248649224),
('use project', 0.04094273528253555),
('hooks', 0.036353940905693174),

```

```

('vite', 0.033015156132700896),
('advanced data', 0.033015156132700896),
('beginners 2023', 0.033015156132700896),
('projects beginners', 0.03174447123621575)],
92: [('reels', 0.1329238126121255),
('instagram stories', 0.12658044132680604),
('instagram', 0.11912391396673522),
('instagram reels', 0.10032643121777804),
('stories', 0.08830402903386071),
('instagram story', 0.059261606761362405),
('story', 0.052114301461032506),
('use instagram', 0.04926280709012626),
('highlights', 0.030015317611990465),
('ux journey', 0.025564833288595823)],
93: [('saas', 0.1834683866774736),
('b2b saas', 0.14774154910003456),
('b2b', 0.08838261850956868),
('saas sales', 0.05866542717875691),
('saas marketing', 0.048651877522714056),
('sales', 0.040978233618464255),
('saas business', 0.03715245486191193),
('marketing saas', 0.0367058333731748),
('saas companies', 0.03395208064877202),
('companies', 0.03039122515360057)],
94: [('sql', 0.2378597418021839),
('queries', 0.0840534337797646),
('sql queries', 0.06641218080033671),
('sql query', 0.055206914391910084),
('advanced', 0.045100475723030393),
('window functions', 0.042046250704399865),
('query', 0.04052537046521834),
('window', 0.03745949994341533),
('abdelilah', 0.03672914027330748),
('abdelilah moulida', 0.03672914027330748)],
95: [('chatgpt', 0.0651735174029379),
('use chatgpt', 0.05269166920866559),
('chatgpt prompts', 0.037156367779161086),
('prompts', 0.03670210269052994),
('ux planet', 0.027690555798069404),
('planet', 0.02717168324343108),
('chat gpt', 0.02616081123477012),
('interfaces', 0.025007445479466234),
('ux', 0.023604582346419867),
('ways use', 0.02279997974171849)],
96: [('money ai', 0.058081151574317695),
('make money', 0.05185241620690745),
('ai', 0.049299844710302056),

```

```

('money', 0.04408068612323988),
('using ai', 0.04308674562678574),
('make', 0.036280719532390544),
('ai make', 0.03423267635661311),
('adventures', 0.028762291047263498),
('money 2023', 0.026809888728760275),
('ways using', 0.023874596376953126)],
97: [('netflix', 0.1411999869049227),
('recommender', 0.12577616154518156),
('recommendation', 0.10018772218722978),
('recommender systems', 0.062334374169101744),
('shopify', 0.05115433750386018),
('start problem', 0.043656443743322254),
('content based', 0.042229736784540596),
('cold start', 0.04102440726977072),
('performant', 0.04102440726977072),
('systems', 0.03724234068313452)],
98: [('blog', 0.15616651895304567),
('blogging', 0.07196537573289828),
('notability', 0.041534822178226224),
('articles', 0.038448629940183734),
('views', 0.029701437666803224),
('visitors', 0.026755820993676047),
('notability blog', 0.02267125271842442),
('software developers', 0.02267125271842442),
('blog ai', 0.02267125271842442),
('generation tool', 0.02267125271842442)],
99: [('days', 0.08054979471152067),
('earned', 0.0629671887806282),
('000', 0.04948859217871662),
('selling', 0.04904811192907873),
('31 days', 0.03895046510037746),
('just', 0.03794328887245467),
('selling ai', 0.036227420264041285),
('400', 0.0275510559292374),
('amazon kdp', 0.0275510559292374),
('43', 0.02725644989743161)],
100: [('advertising', 0.10353575936118757),
('ad', 0.07326957392439515),
('ads', 0.05188489956964495),
('hopkins', 0.0375172228780692),
('long lasting', 0.0375172228780692),
('don suck', 0.0375172228780692),
('eu', 0.03489437666125189),
('psychological triggers', 0.03389841444856055),
('lasting', 0.03389841444856055),
('claude', 0.03303633983917133)],

```



```

101: [('christian', 0.15385639034391932),
      ('church', 0.1488124792007545),
      ('christians', 0.14458953769309785),
      ('jesus', 0.14197614829277658),
      ('bible', 0.12561652148320823),
      ('backyard', 0.1054366431783684),
      ('church medium', 0.1054366431783684),
      ('backyard church', 0.1054366431783684),
      ('god', 0.09420012311775844),
      ('gay', 0.05411362880027436)],
102: [('ux writing', 0.2561829885029934),
      ('writing', 0.1161037424205202),
      ('ux', 0.10625919013788683),
      ('ux writer', 0.07467281136279237),
      ('writing challenge', 0.07420907108908481),
      ('ux writers', 0.05542043938161638),
      ('challenge', 0.049272664026376686),
      ('writer', 0.0429182360598857),
      ('terms', 0.03987516879951493),
      ('design ux', 0.03926899562135995)],
103: [('color', 0.2696799589517362),
      ('dark mode', 0.08199511767138729),
      ('colors', 0.07990989008218403),
      ('contrast', 0.07807200953372079),
      ('mode', 0.06490831929088707),
      ('buttons', 0.062453596123982054),
      ('dark', 0.060231302740367),
      ('ui design', 0.04492782333378991),
      ('design', 0.03953382932905376),
      ('theory', 0.03803465028927269)],
104: [('snowflake', 0.3129637845897002),
      ('snowflake data', 0.0626442699009399),
      ('medium snowflake', 0.052203558250783245),
      ('stored', 0.042989872046662325),
      ('data warehouse', 0.03976467685584179),
      ('snowflake medium', 0.03892795843415992),
      ('elt', 0.03817163795097783),
      ('warehouse', 0.03817163795097783),
      ('sql server', 0.03466591393933594),
      ('data snowflake', 0.03466591393933594)],
105: [('chatgpt', 0.07112723748446366),
      ('chatgpt ai', 0.060144564314894136),
      ('ai tools', 0.049627825873528916),
      ('forget', 0.04269001190312548),
      ('ai', 0.04212895820698539),
      ('blow mind', 0.03377471043666844),
      ('tools blow', 0.03286336930234859),

```

```

('blow', 0.03241929471919896),
('tools', 0.030534941641732573),
('power artificial', 0.028649515652343752)],
106: [('autistic', 0.30083306603656346),
('nuance medium', 0.04647172492090829),
('exhausting', 0.04647172492090829),
('nuance', 0.04647172492090829),
('obviously', 0.04647172492090829),
('free tool', 0.04427319787910081),
('laziness', 0.04117803835887196),
('time management', 0.04117803835887196),
('explains', 0.04000272662895268),
('forces', 0.04000272662895268)],
107: [('elasticsearch', 0.3199684980427833),
('search', 0.13380303991735298),
('action', 0.10627163294201128),
('db', 0.07489331904147917),
('mongodb', 0.06141733730559967),
('elastic', 0.06021986413374262),
('queries', 0.059061936628057594),
('autocomplete', 0.05825185914181887),
('fastapi angular', 0.046972783115631356),
('project using', 0.04244191727705955)],
108: [('clients', 0.20597072742975606),
('client', 0.13320486875611148),
('consulting', 0.08671423630744345),
('consultant', 0.05898931662529246),
('difficult', 0.05218224991174252),
('consultants', 0.04860696507402955),
('ve known', 0.046469053537983834),
('morons', 0.046469053537983834),
('years building', 0.0446805560295262),
('honestly', 0.04322038074396347)],
109: [('jobs pay', 0.16363722986772425),
('experience required', 0.16080119037483503),
('media jobs', 0.15793915968258512),
('required', 0.15479637692478626),
('work home', 0.15207941115963758),
('50 hour', 0.15107479601027562),
('hour', 0.12631564809433163),
('jobs', 0.12159745911455465),
('25', 0.12033135128238549),
('home', 0.1199049187123093)],
110: [('america', 0.16847717503681003),
('eudaimonia', 0.1641820254338645),
('american', 0.15263943797349078),
('fascism', 0.15206233763343718),

```

```

('collapse', 0.0985713147901793),
('trump', 0.09377939388647083),
('republicans', 0.0677068631627655),
('collapse eudaimonia', 0.06510096640239559),
('american fascism', 0.06510096640239559),
('world', 0.0581777983720744)],
111: [('money online', 0.1464355276125809),
('online', 0.09696570735123432),
('money', 0.09162531505221681),
('make money', 0.09005945972778662),
('ways make', 0.06369412338501468),
('make', 0.046117013930232),
('making money', 0.04032540975568601),
('ways', 0.039024062047673364),
('earn money', 0.032824681018309536),
('way make', 0.03084409496901454)],
112: [('ai ux', 0.06336863762744961),
('ai design', 0.061911196172351544),
('designers', 0.059875185310812575),
('10 ai', 0.05090245505312841),
('ux', 0.050243739558549404),
('ai', 0.04824717898695041),
('design', 0.04425935090051905),
('design opportunities', 0.03953937698461037),
('opportunities challenges', 0.03953937698461037),
('ux design', 0.039446749843826144)],
113: [('speech', 0.19448191945502813),
('audio', 0.180597525910463),
('whisper', 0.16675802010248256),
('speech recognition', 0.12265950498386193),
('recognition', 0.08710860746551129),
('openai whisper', 0.08156348702332336),
('audio classification', 0.057123727160285416),
('text', 0.055736679072265156),
('sound', 0.049063801993544764),
('automatic', 0.044634909775841405)],
114: [('agile', 0.2990209162185333),
('scrum', 0.27579853880444544),
('2023 ux', 0.050436710833774406),
('sprints', 0.04966463315091109),
('sprint', 0.04966463315091109),
('kanban', 0.04966463315091109),
('master', 0.048409973149992847),
('safe', 0.037827533125330805),
('coaching', 0.03637006361998093),
('practices', 0.03246943524546656)],
115: [('climate', 0.33841408336494144),

```

```

('climate medium', 0.11028920757207727),
('new climate', 0.11028920757207727),
('change', 0.10429696492392537),
('oil', 0.09190767297673107),
('environment', 0.07542874211162345),
('heat', 0.05676479583626187),
('sustainability', 0.05514460378603864),
('civilization', 0.05040269095639678),
('air', 0.0464852431716213)],
116: [('scraping', 0.24988880470730707),
('web scraping', 0.23441512176451051),
('web', 0.15203699743623045),
('scrape', 0.11854236971541199),
('medium web', 0.04284986730155366),
('python', 0.03836041957336951),
('guide scrape', 0.03735551071049131),
('zillow', 0.03591777456683894),
('scraping data', 0.026140345268010914),
('scrape amazon', 0.026140345268010914)],
117: [('tiktok', 0.36938327796756),
('account', 0.03512755238986937),
('likes', 0.0316706231701754),
('free week', 0.030629091425144105),
('tiktok instagram', 0.029180062238498266),
('content calendar', 0.029180062238498266),
('media platform', 0.028056982153220996),
('marketing tiktok', 0.028056982153220996),
('year review', 0.02714007073652925),
('yard couch', 0.02636543345999154)],
118: [('tesla', 0.3208009265387172),
('ev', 0.14001646780225982),
('predict medium', 0.12201171977250184),
('predict', 0.11023764059183251),
('musk', 0.09369674206165353),
('battery', 0.09190767297673107),
('trouble', 0.08566892261038715),
('embarrassed', 0.07824341501884165),
('mile', 0.06103153862559144),
('just', 0.05284746024488992)],
119: [('old', 0.09060888713325364),
('year old', 0.07504498698303322),
('wish', 0.07130405956685767),
('wish knew', 0.07103912527206821),
('age', 0.06971326664662458),
('knew', 0.06049007505464017),
('aging', 0.05287855964872579),
('year', 0.047933500323469144),

```

```

('40', 0.045644756684861024),
('regrets', 0.04432406223141827)],
120: [('chatgpt prompts', 0.1330116855919813),
('prompts', 0.12766203958404387),
('chatgpt', 0.08770938542680795),
('awesome', 0.06261715326843789),
('prompt', 0.06165948397663923),
('chatgpt prompt', 0.05344306717596938),
('prompts chatgpt', 0.0434125224538738),
('medium awesome', 0.04230849343228594),
('use chatgpt', 0.03665580698187995),
('secret chatgpt', 0.036035253575193285)],
121: [('pr', 0.23409651781385032),
('leverage', 0.06378430343126949),
('superstar medium', 0.06021986413374262),
('aka pr', 0.06021986413374262),
('jill', 0.06021986413374262),
('pr superstar', 0.06021986413374262),
('jill kent', 0.06021986413374262),
('kent aka', 0.06021986413374262),
('aka', 0.05825185914181887),
('kent', 0.05825185914181887)],
122: [('cybersecurity', 0.3450411486547568),
('security', 0.11596235482099829),
('cyber', 0.08219585818023632),
('cyber security', 0.07507138468023997),
('threat', 0.07433451047714218),
('complete list', 0.07261802710584853),
('ai replace', 0.06875130182746465),
('jobs', 0.05507428983251018),
('replace', 0.04800454068945165),
('resources', 0.04249256080619262)],
123: [('rust', 0.4365678579942992),
('head', 0.051771173479419855),
('coding', 0.0416324405595996),
('level coding', 0.041021861134574344),
('testing using', 0.03755145574104355),
('head head', 0.03755145574104355),
('world write', 0.0357749369858988),
('haskell', 0.0357749369858988),
('hard understand', 0.0357749369858988),
('explained level', 0.0357749369858988)],
124: [('android', 0.350888540130032),
('android apps', 0.07982195451190098),
('android interview', 0.053005984616721624),
('activity', 0.050965894186584926),
('architecture', 0.04843630404000443),

```

```

('architect', 0.03982334840092273),
('interview questions', 0.03974533546622055),
('context', 0.037912810350334655),
('testing', 0.037887147521482346),
('app architecture', 0.037092110716688276)],
125: [('peace', 0.1471546045064207),
('mindfulness', 0.12437047423054287),
('peace mind', 0.11378010973363158),
('quotes', 0.07240574270212868),
('marcus aurelius', 0.07099402512891702),
('aurelius', 0.07099402512891702),
('marcus', 0.06671381246106485),
('buddhist', 0.05537685932801268),
('bring', 0.054070935903613816),
('mind', 0.05222883198312862)],
126: [('father', 0.10914547680215565),
('car', 0.0990645084943875),
('daughter', 0.09792645327456105),
('left', 0.09481644879516797),
('mother', 0.0834237539383283),
('child', 0.0834237539383283),
('mom', 0.0834237539383283),
('grace period', 0.08214671549605672),
('ago', 0.08121995359128364),
('year ago', 0.07422306860300935)],
127: [('note', 0.28121102730004194),
('note taking', 0.2657336383273695),
('taking', 0.22433598881429215),
('notes', 0.15295635600974464),
('according research', 0.06735943000101063),
('visual', 0.059626944897820326),
('10 apps', 0.05591276441828377),
('tana', 0.05591276441828377),
('apps', 0.05519218678423043),
('dialog', 0.04923480324418436)],
128: [('pdfs', 0.10499076512407007),
('extracting', 0.10004200282403664),
('invoice', 0.09975679088690949),
('pdf', 0.08296177910701533),
('text images', 0.07311422724199335),
('text', 0.060732683919078836),
('recognition', 0.05694999714972082),
('python', 0.05510598972768871),
('document', 0.04975260519635045),
('optical', 0.04855170019514837)],
129: [('web3', 0.317212102610045),
('nft', 0.05259773626901664),

```

```

('dapps', 0.04694604901130498),
('web3 marketing', 0.04694604901130498),
('decentralized', 0.03620579621357925),
('pitch deck', 0.03416663043240464),
('web3 community', 0.03416653578692131),
('future web3', 0.03416653578692131),
('medium web3', 0.03416653578692131),
('web', 0.03406607652816624)],
130: [('reality', 0.2329925589258736),
('ar', 0.20547980659891274),
('augmented reality', 0.18520255988214174),
('augmented', 0.15037218985629414),
('vr', 0.11459561053556712),
('virtual reality', 0.07743132573246096),
('immersive', 0.07330830454225336),
('app design', 0.06194506058596877),
('mixed', 0.05864664363380269),
('virtual', 0.056523554193416044)],
131: [('docker', 0.25205935797714957),
('kubernetes', 0.22806874830077975),
('eks', 0.0651361697676702),
('docker image', 0.0651361697676702),
('docker compose', 0.0651361697676702),
('container', 0.06024951038763908),
('cluster', 0.05781539199631526),
('supervisor', 0.052524112029296875),
('creating multi', 0.052524112029296875),
('configuring', 0.052524112029296875)],
132: [('ambassador', 0.4943983809285635),
('ambassador program', 0.4364307739051803),
('program', 0.3422648188173523),
('sei', 0.07624531970823759),
('ambassador programs', 0.0639592508105829),
('july', 0.061705125586172),
('join', 0.05403011801912133),
('phase', 0.052820532219987774),
('protocol', 0.049421118553756097),
('dojima network', 0.0447569011230888)],
133: [('javascript', 0.26022379353008074),
('developer', 0.06505594838252018),
('developer know', 0.06232707971135077),
('javascript plain', 0.06090525201579209),
('shorthand', 0.05433528830616918),
('useful javascript', 0.052244035733583924),
('dom', 0.04909425540826011),
('plain english', 0.04781280835804447),
('plain', 0.04585447588562813),

```

```

('english', 0.043434704961266245)],
134: [('swiftui', 0.30420938468838354),
      ('swift', 0.22135073565309116),
      ('ios', 0.08655628402349777),
      ('networking', 0.053213249931660545),
      ('writing clean', 0.045089666187200984),
      ('efficient code', 0.045089666187200984),
      ('devtechie medium', 0.045089666187200984),
      ('widgets', 0.045089666187200984),
      ('devtechie', 0.045089666187200984),
      ('building large', 0.045089666187200984)],
135: [('airflow', 0.378010543082015),
      ('apache airflow', 0.18520255988214174),
      ('mage', 0.10263162635915471),
      ('apache', 0.09719817889295834),
      ('orchestration', 0.09291519840736824),
      ('dags', 0.06403783822867731),
      ('datageeks', 0.06194506058596877),
      ('datageeks medium', 0.06194506058596877),
      ('skip', 0.05609215912703158),
      ('dag', 0.04802837867150799)],
136: [('scala', 0.35404904828466655),
      ('functional', 0.14495001490576534),
      ('functional programming', 0.10224545879726205),
      ('medium functional', 0.09595953581844269),
      ('medium scala', 0.07457651178683321),
      ('programming', 0.06924330293374707),
      ('anti patterns', 0.06190341774881418),
      ('zio', 0.05213565861717846),
      ('telemetry', 0.043318286444132374),
      ('softwaremill', 0.043318286444132374)],
137: [('topic', 0.31784585658036196),
      ('topic modeling', 0.26915925483079434),
      ('bertopic', 0.22872158252989894),
      ('modeling', 0.2198247370581758),
      ('topic modelling', 0.09278291188873647),
      ('modelling', 0.07453021208821059),
      ('modelling python', 0.06754142816216117),
      ('latent', 0.06533415204659929),
      ('unsupervised', 0.05082701833997754),
      ('topics', 0.05024311645211522)],
138: [('exam', 0.22026050235691058),
      ('associate', 0.15755024976201154),
      ('certification', 0.14364292049846042),
      ('certified', 0.12527307202303817),
      ('passed', 0.12315345243451788),
      ('aws', 0.1022053578056259),

```



```

('data engineer', 0.09200287448543072),
('specialty', 0.08161919254383504),
('certifications', 0.07806408776903728),
('engineer', 0.07734859006263359)],
139: [('sem', 0.2326035740036531),
('engine marketing', 0.13418553853951448),
('search engine', 0.09417471497744898),
('engine', 0.08891968342085146),
('vs sem', 0.08708715533369109),
('seo vs', 0.08385168489774725),
('seo sem', 0.07337022428552885),
('seo', 0.0659422410790405),
('marketing sem', 0.06531536650026831),
('search', 0.05713182333892994)],
140: [('clustering', 0.4131469979452329),
('unsupervised', 0.1254156677538946),
('means', 0.11517548676441437),
('kmeans', 0.07407043288450343),
('scikit learn', 0.06487993072359986),
('methods', 0.06461757939459113),
('scikit', 0.06141252161401292),
('evaluation', 0.06141252161401292),
('plot python', 0.05777652323222657),
('scatter', 0.05777652323222657)],
141: [('github', 0.1913646874742477),
('github copilot', 0.08639443739281014),
('room database', 0.08639443739281014),
('copilot', 0.07991293580354057),
('commit', 0.07823142772138247),
('flow', 0.07642175014986358),
('room', 0.06888026610836354),
('coding use', 0.06714284019393343),
('like senior', 0.06714284019393343),
('senior engineer', 0.0654353226396448)],
142: [('obsidian', 0.4929487070496144),
('second brain', 0.0722994124651669),
('note', 0.06604198368410075),
('fusing', 0.06302893443515625),
('obsidian vault', 0.06302893443515625),
('notes', 0.057474509530934345),
('note taking', 0.05705795870579016),
('building second', 0.05694933627358172),
('vault', 0.055501050929807824),
('second', 0.053667443535357746)],
143: [('neural', 0.19394063703956138),
('neural networks', 0.1420994283719647),
('networks', 0.12550396104262918),

```

```

('neural network', 0.10325817133306035),
('deep learning', 0.08768562487988554),
('deep', 0.073451201256119),
('recurrent neural', 0.06735943000101063),
('network', 0.06694466946459335),
('recurrent', 0.06564640432557914),
('learning', 0.05728760824673902)],
144: [('writing online', 0.07737494179252263),
('online', 0.07383967378519835),
('write', 0.058456077630383574),
('online writers', 0.05809442168467882),
('writing', 0.05694264626695909),
('write articles', 0.053736049805191616),
('articles', 0.04765254907644394),
('read online', 0.04684582964775127),
('frustration', 0.045042830808143974),
('things', 0.041274800242312046)],
145: [('fintech', 0.19265079532088264),
('markets', 0.10612762951730342),
('economics', 0.09905500471581181),
('finance', 0.09684469703011363),
('defi', 0.07162682172548453),
('asset', 0.06727181757055285),
('embedded', 0.06596481212726585),
('banking', 0.06171946141966776),
('decentralized', 0.059230680441569754),
('olive', 0.05589456315371919)],
146: [('career', 0.22058765452473844),
('professional development', 0.07568976842649768),
('brutal truth', 0.07568976842649768),
('transformed', 0.07390481632240666),
('brutal', 0.06664973364933253),
('harder', 0.05937903266139635),
('passing', 0.05937903266139635),
('banish', 0.05296559041203958),
('layoff', 0.05045984561766512),
('pursuing', 0.04851775079770967)],
147: [('week scala', 0.3451118311909425),
('petr', 0.3157176001368623),
('medium week', 0.28377414864262046),
('scala', 0.2644573683833637),
('week', 0.2331009908065509),
('2023 week', 0.0931984613808792),
('scala jul', 0.08396459384113736),
('2023', 0.08040008034040745),
('jul', 0.07455398037300358),
('scala sep', 0.06717167507290989)],

```

```

148: [('weekly', 0.10326066364739621),
      ('credix', 0.1029450958971799),
      ('presearch', 0.1029450958971799),
      ('2023', 0.07949581543977995),
      ('updates', 0.07185137831862391),
      ('august', 0.07058910530346311),
      ('news', 0.06296759071379363),
      ('update', 0.06155042664688414),
      ('medium 2023', 0.06000408994342901),
      ('trend', 0.050581643910882955)],
149: [('keywords', 0.23163139265807864),
      ('keyword', 0.2014757737593105),
      ('keyword research', 0.1576335410237666),
      ('research', 0.09899324775598464),
      ('volume', 0.0427294454675125),
      ('seo', 0.03999481921618874),
      ('chatgpt keyword', 0.03620632896823005),
      ('good seo', 0.0344934467058069),
      ('seo keywords', 0.0344934467058069),
      ('nolan medium', 0.03358783834237763)],
150: [('spotify', 0.47397333039099404),
      ('duolingo', 0.11860787436182751),
      ('critique', 0.11671930025287353),
      ('freemium', 0.09536965607446793),
      ('yip medium', 0.08961984249061501),
      ('jason yip', 0.08961984249061501),
      ('yip', 0.08961984249061501),
      ('playlists', 0.08961984249061501),
      ('bye', 0.08043283937616527),
      ('premium', 0.07781286683524902)],
151: [('illumination medium', 0.27092590994310006),
      ('illumination', 0.23509826126522257),
      ('kdp', 0.0678673816785999),
      ('amazon kdp', 0.04242290618861814),
      ('declutter', 0.04196927406351926),
      ('discover secret', 0.04196927406351926),
      ('medium stand', 0.04196927406351926),
      ('stand crowd', 0.04196927406351926),
      ('amazon search', 0.04196927406351926),
      ('amazon', 0.041707573072988696)],
152: [('pinterest', 0.41341572305969815),
      ('affiliate', 0.058589493001990296),
      ('using free', 0.05682936711366548),
      ('drove', 0.05464212261971565),
      ('sales using', 0.05464212261971565),
      ('affiliate marketing', 0.05452220195430343),
      ('free traffic', 0.046932341743005535),

```

```

('month using', 0.04608406124913077),
('traffic', 0.0423812233380702),
('slump', 0.03976760722740021)],
153: [('reinforcement learning', 0.24778025141277965),
('reinforcement', 0.24778025141277965),
('policy', 0.14494722000073487),
('learning', 0.131793384463741),
('ppo', 0.10169524334568164),
('learning deep', 0.08723594165312971),
('deep', 0.08064889866374339),
('deep reinforcement', 0.072146525536854),
('gradient', 0.06607267967834266),
('optimization', 0.0633732362722181)],
154: [('knowledge', 0.20858124233762348),
('knowledge graph', 0.17313512723592045),
('graph', 0.16013725282863925),
('knowledge graphs', 0.15496066643258993),
('graphs', 0.13162897783624133),
('augmented generation', 0.12476909754538436),
('retrieval augmented', 0.12476909754538436),
('retrieval', 0.12054275596745816),
('augmented', 0.08919512800022916),
('question answering', 0.058867469062485425)],
155: [('britain', 0.6105305394810892),
('eudaimonia', 0.28020399007379543),
('collapse', 0.14419575192163372),
('cares', 0.11932241885893313),
('british', 0.09212115438570499),
('portugal', 0.0693092583106118),
('collapsing', 0.06603031226540179),
('destroyed', 0.061414102923803324),
('mistake', 0.04527078913639537),
('global', 0.043609324248649224)],
156: [('amazon', 0.32525286747237053),
('deceptive patterns', 0.06878157527646019),
('sued', 0.06878157527646019),
('deceptive', 0.06613431507544948),
('amazon kdp', 0.0648688882989981),
('kdp', 0.062265629540048786),
('designers learn', 0.06214709315569433),
('shoppers', 0.0481314293823693),
('want think', 0.0481314293823693),
('amazon prime', 0.0481314293823693)],
157: [('kotlin', 0.544048008228886),
('multiplatform', 0.18449269400075793),
('compose multiplatform', 0.1053134116841281),
('kotlin multiplatform', 0.1053134116841281),

```

```

('failure', 0.08121995359128364),
('coroutines', 0.07898505876309607),
('scope', 0.06918476025028421),
('strategic', 0.06885894033413349),
('compose', 0.05992602889481307),
('reusability', 0.05748398201117093)],
158: [('delta lake', 0.2823497544354979),
('lake', 0.26962481130089494),
('delta', 0.2503840019169417),
('databricks', 0.10967921064654701),
('loader', 0.08614448556234858),
('using auto', 0.07333706226188458),
('incremental data', 0.06561827864000296),
('keeping', 0.06298950156927036),
('auto', 0.060885094769898424),
('merge', 0.06080447684555734)],
159: [('prompt engineering', 0.2790845238653442),
('prompt', 0.22430580347468168),
('engineering', 0.16912510928048555),
('advanced prompt', 0.08714168574647461),
('engineering technique', 0.04756517727198849),
('guide prompt', 0.04531492018213848),
('introduced', 0.04357084287323731),
('advanced', 0.04280123578420923),
('prompting', 0.04241051527106875),
('science best', 0.042146933379080716)],
160: [('jetpack', 0.4483541767478186),
('jetpack compose', 0.44206848137130583),
('compose', 0.4355290611121914),
('navigation', 0.15543020314430142),
('compose navigation', 0.07407043288450343),
('animations', 0.06627446142640299),
('screen', 0.0578765996760077),
('navigation bar', 0.04938028858966895),
('android', 0.047437954572618284),
('bar', 0.03760933986184695)],
161: [('influencer', 0.3275272617081631),
('influencer marketing', 0.3054392472638236),
('influencers', 0.1276495346095619),
('medium influencer', 0.06615632431170981),
('marketing', 0.05141090315978468),
('increasing user', 0.04629435192502696),
('hotel booking', 0.04629435192502696),
('twitch', 0.04410421620780654),
('marketing brand', 0.04102087027353275),
('marketing vs', 0.03883661324604873)],
162: [('emotionally', 0.1861793585489038),

```

```

('habits emotionally', 0.14972291705446122),
('habits', 0.13907871482415846),
('people', 0.11281950981138816),
('nick wignall', 0.10772607297727932),
('wignall medium', 0.10772607297727932),
('wignall', 0.10772607297727932),
('people nick', 0.10334103857370752),
('nick', 0.08421801587512505),
('confident', 0.07486145852723061)],
163: [('podcast', 0.47700242480102456),
('podcasts', 0.14067366507415682),
('pacific content', 0.0635032471432307),
('episodes', 0.0635032471432307),
('advertising best', 0.0635032471432307),
('pacific', 0.06049897720651996),
('standards', 0.05327326005478936),
('host', 0.052048006355813865),
('werner medium', 0.04486233038680851),
('werner', 0.04486233038680851)],
164: [('money instagram', 0.22131649995625316),
('instagram', 0.10963712262820789),
('money', 0.08023438390431764),
('make money', 0.06844518939311783),
('medium make', 0.05285415163275958),
('make', 0.039215586670742736),
('5k month', 0.03555478352752404),
('hustle club', 0.034621310291373865),
('hustle', 0.03149827761474582),
('club', 0.03089112894576269)],
165: [('idor', 0.12076701013777733),
('google dorks', 0.11682030447462588),
('dorks', 0.11682030447462588),
('injection', 0.11229609587463252),
('sql injection', 0.09057525760333299),
('gray area', 0.08294994024835407),
('vulnerabilities', 0.08294994024835407),
('vulnerability', 0.08294994024835407),
('gray', 0.07933687180874983),
('area', 0.07779545778052548)],
166: [('market fit', 0.25238022581101194),
('product market', 0.25238022581101194),
('fit', 0.21494436498756567),
('metrics', 0.16174381025259332),
('market', 0.13369148473865186),
('product', 0.1287328064700295),
('metric', 0.07727946546954852),
('cheat sheet', 0.0668591871378168),

```

```

('sheet', 0.06269380278906939),
('cheat', 0.055716112753301135)],
167: [('transformers', 0.2813166504533626),
('attention', 0.23132757851089572),
('transformer', 0.15236697138208832),
('transformer models', 0.08969575676073714),
('transformer architecture', 0.07217081785822568),
('science visualizing', 0.05589456315371919),
('nikolaiev dimid', 0.05589456315371919),
('nikolaiev', 0.05589456315371919),
('dimid', 0.05589456315371919),
('dimid data', 0.05589456315371919)]]}

```

```

[87]: # Save original representations
from copy import deepcopy
original_topics = deepcopy(topic_model.topic_representations_)

```

```

[88]: def compare_topic_changes(new_model, original_topic_words, max_length=75,
    ↪ top_n_topics=10):

    """Displays differences in top words of topic representations between the
    ↪ original and new models."""

    for topic_id in range(top_n_topics):
        # Extract top 5 words per topic from the original and new models
        original_top_words = "_".join(word for word, _ in
    ↪ original_topic_words[topic_id][:5])
        new_top_words = "_".join(word for word, _ in new_model.
    ↪ get_topic(topic_id)[:5])

        # Calculate whitespace for alignment
        alignment_spaces = " " * (max_length - len(original_top_words))

        # Print 'before' and 'after' topic word changes
        print(f"Topic: {topic_id}    {original_top_words}{alignment_spaces} >> 
    ↪ {new_top_words}")

```

## 0.5 KeyBERTInspired

```

[89]: # KeyBERTInspired
from bertopic.representation import KeyBERTInspired
representation_model = KeyBERTInspired()

# Update our topic representations
topic_model.update_topics(medium_titles,
    ↪ representation_model=representation_model)

```

### 0.5.1 Comparison of Topics

```
[90]: # Show topic differences
```

```
compare_topic_changes(topic_model, original_topics)
```

```
Topic: 0    ux_design_collective_ux collective_ui
>> designs_design_designing_designers_ux
Topic: 1    email_email marketing_newsletter_emails_substack
>> newsletters_emails_newsletter_marketing_mail
Topic: 2    startup_founders_venture_entrepreneur_startups
>> startups_startup_entrepreneurs_entrepreneurship_founders
Topic: 3    b2b_sales_marketing_b2b sales_funnel
>> sales_b2b_marketing_b2c_business
Topic: 4    data_data engineering_data analyst_data science_analyst
>> data_analyst_analytics_analytical_study
Topic: 5    brand_branding_identity_brand identity_logo
>> branding_brands_rebranding_rebrands_brand
Topic: 6    medium_jessica_holiday medium_ryan holiday_holiday
>> song_jessica_ryan_chris_music
Topic: 7    linkedin_profile_leads_recruiters_using linkedin
>> linkedin_jobs_recruiters_company_recruiter
Topic: 8    google ads_ads_google_adwords_google adwords
>> adwords_ads_adsense_advertising_ad
Topic: 9    social media_social_media_media marketing_medium social
>> marketing_advertising_social_marketer_businesses
```

```
[91]: topic_model.get_topic(0, full=True)
```

```
[91]: {'Main': [('designs', 0.53380334),
                ('design', 0.51747537),
                ('designing', 0.50500005),
                ('designers', 0.48914835),
                ('ux', 0.47295764),
                ('trends', 0.43599612),
                ('innovation', 0.39386082),
                ('industry', 0.36847156),
                ('designer', 0.36549464),
                ('future', 0.36333477)]}
```

```
[92]: # topic_model.topic_representations_
```

### 0.5.2 Hierarchical Topics

```
[93]: hierarchical_topics = topic_model.hierarchical_topics(medium_titles)
      topic_model.visualize_hierarchy(hierarchical_topics=hierarchical_topics)
```

```
100%|          | 167/167 [00:28<00:00, 5.95it/s]
```



Above you can see 167 meaningful clusters/topics, with around 10k documents as outliers. In the following code, I am attempting to remove outliers based on the probabilities strategy

### 0.5.3 Remove Outliers

```
[95]: new_topics = topic_model.reduce_outliers(medium_titles, topics,
↳ probabilities=probs, strategy="probabilities")
```

```
[96]: # prompt: get values counts from elements of a list
```

```
from collections import Counter
c = Counter(new_topics)
```

```
[97]: print(c)
print(len(c))
```

```
Counter({0: 972, 2: 695, 3: 648, 6: 616, 1: 599, 4: 510, 125: 489, 5: 435, 126:
418, 12: 417, 10: 405, 8: 388, 7: 383, 39: 337, 26: 329, 11: 327, 25: 323, 19:
306, 9: 302, 14: 302, 21: 300, 16: 296, 13: 293, 110: 287, 15: 273, 83: 270,
119: 258, 35: 250, 20: 249, 38: 247, 17: 244, 68: 242, 18: 234, 32: 233, 99:
232, 111: 229, 28: 226, 31: 221, 162: 216, 146: 209, 24: 208, 41: 203, 50: 202,
33: 200, 42: 199, 30: 194, 144: 187, 46: 187, 65: 186, 34: 184, 22: 184, 55:
182, 64: 181, 23: 180, 40: 172, 95: 171, 27: 170, 36: 169, 82: 168, 29: 168, 37:
165, 120: 162, 44: 162, 72: 161, 59: 160, 43: 156, 56: 155, 54: 155, 87: 154,
58: 151, 98: 150, 62: 144, 135: 144, 57: 141, 88: 141, 47: 141, 48: 138, 49:
138, 104: 138, 149: 136, 115: 135, 84: 133, 61: 131, 101: 131, 60: 131, 143:
131, 51: 129, 45: 128, 86: 125, 105: 123, 70: 123, 52: 123, 92: 121, 93: 121,
165: 120, 157: 119, 91: 116, 74: 114, 53: 113, 77: 109, 166: 109, 141: 109, 89:
107, 148: 106, 63: 105, 113: 105, 80: 104, 158: 102, 151: 101, 69: 99, 75: 98,
112: 98, 145: 98, 97: 97, 66: 96, 129: 96, 96: 94, 123: 94, 67: 92, 100: 91,
133: 91, 102: 91, 94: 91, 85: 91, 73: 90, 78: 89, 108: 85, 71: 84, 76: 80, 79:
78, 131: 78, 124: 77, 118: 77, 137: 77, 81: 75, 164: 74, 155: 74, 90: 69, 154:
69, 167: 69, 121: 65, 136: 63, 103: 61, 122: 60, 140: 60, 128: 60, 134: 59, 106:
58, 139: 57, 107: 57, 109: 56, 114: 53, 160: 53, 117: 52, 116: 52, 130: 50, 127:
49, 156: 48, 132: 46, 138: 42, 142: 41, 159: 41, 147: 40, 153: 38, 150: 38, 152:
37, 161: 33, 163: 31})
168
```

```
[100]: topic_model.update_topics(medium_titles, topics=new_topics)
```

2023-12-13 23:44:28,976 - BERTopic - WARNING: Using a custom list of topic assignments may lead to errors if topic reduction techniques are used afterwards. Make sure that manually assigning topics is the last step in the pipeline. Note that topic embeddings will also be created through weightedc-TF-IDF embeddings instead of centroid embeddings.

```
[101]: topic_model.get_topic_info().head(20)
```

```
[101]:
```

	Topic	Count	Name \
0	0	972	0_ux_design_collective_ui
1	1	599	1_email_newsletter_emails_substack
2	2	695	2_startup_entrepreneur_founders_startups
3	3	648	3_b2b_sales_marketing_customers
4	4	510	4_data_science_engineering_analyst
5	5	435	5_brand_branding_identity_logo
6	6	616	6_medium_the_jessica_poetry
7	7	383	7_linkedin_410_deleted_author
8	8	388	8_ads_google_adwords_ad
9	9	302	9_social_media_marketing_management
10	10	405	10_women_men_woman_relationship
11	11	327	11_midjourney_prompts_images_v5
12	12	417	12_seo_website_traffic_ranking
13	13	293	13_product_manager_management_managers
14	14	302	14_image_opencv_computer_vision
15	15	273	15_object_detection_yolov8_yolo
16	16	296	16_digital_marketing_agency_company
17	17	244	17_copywriting_copywriter_copy_handbook
18	18	234	18_diffusion_stable_controlnet_sd-xl
19	19	306	19_ai_art_dall_image

	Representation \
0	[ux, design, collective, ui, designers, design...
1	[email, newsletter, emails, substack, list, su...
2	[startup, entrepreneur, founders, startups, in...
3	[b2b, sales, marketing, customers, funnel, cus...
4	[data, science, engineering, analyst, engineer...
5	[brand, branding, identity, logo, brands, coff...
6	[medium, the, jessica, poetry, thief, taylor, ...
7	[linkedin, 410, deleted, author, profile, lead...
8	[ads, google, adwords, ad, campaign, advertisi...
9	[social, media, marketing, management, strateg...
10	[women, men, woman, relationship, dating, tran...
11	[midjourney, prompts, images, v5, create, cons...
12	[seo, website, traffic, ranking, search, rank,...
13	[product, manager, management, managers, coal...
14	[image, opencv, computer, vision, python, pyto...
15	[object, detection, yolov8, yolo, yolov7, mode...
16	[digital, marketing, agency, company, presence...
17	[copywriting, copywriter, copy, handbook, watk...
18	[diffusion, stable, controlnet, sd-xl, stabledi...
19	[ai, art, dall, image, images, mlearning, gene...

	Representative_Docs
0	[Make UX simple again - UX Collective, What th...
1	[The Best 3 Email Scraping Tools That Help Ema...

```

2  [Is Your Startup a Good Fit for Venture Capita...
3  [B2B sales needs a major overhaul - Sandhya Do...
4  [Becoming a "real" data analyst - Towards Data...
5  [The Brand Archetype Strategy - Umair Asghar -...
6  [If You Try To Do Everything, You Won't Do Any...
7  [5 Tools for LinkedIn Marketing - Sangita Cham...
8  [Digital Marketing (Google Ads Definition - Co...
9  [Developing a Social Media Marketing Strategy ...
10 [Meet the Female Dating Strategy: Women Teachi...
11 [The Ultimate Guide to MidJourney V5.1 (with i...
12 [SEO Guide: To Show You Path for SEO - Chandni...
13 [The Five Whys of a Product Manager - Kubo - M...
14 [Introduction to Image Processing with Python,...
15 [DEtection TRansformer (DETR) vs. YOLO for obj...
16 [Digital Marketing - Moose - Medium, What is D...
17 [on copywriting, on copywriting, What About Co...
18 [Stable Diffusion as an API, Stable Diffusion,...
19 [5 AI Art Generators You Should Know About - M...

```

The outliers have been reduced to 972

```
[102]: hierarchical_topics = topic_model.hierarchical_topics(
medium_titles)
topic_model.visualize_hierarchy(hierarchical_topics=hierarchical_topics)
```

100%| | 167/167 [00:01<00:00, 130.17it/s]

[ ]:

#### 0.5.4 Merge Topics

```
[103]: topics_to_merge = [[105, 25, 95, 120 ], [44, 62, 163], [124, 160], [142, 127]]

topic_model.merge_topics(medium_titles, topics_to_merge)
```

In the 1st list merged topics similar to ChatGPT, 2nd list related to podcasts, 3rd list about Android and 4th list about note taking

```
[104]: hierarchical_topics = topic_model.hierarchical_topics(medium_titles)
topic_model.visualize_hierarchy(hierarchical_topics=hierarchical_topics)
```

100%| | 160/160 [00:01<00:00, 121.96it/s]

[ ]: