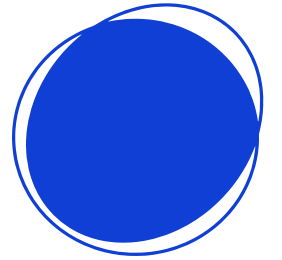


Generowanie tytułów publikacji naukowych

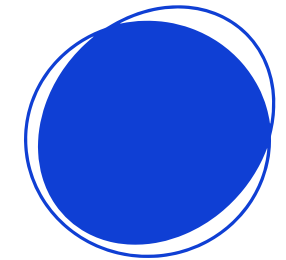
Stanisław Greń

Cel projektu



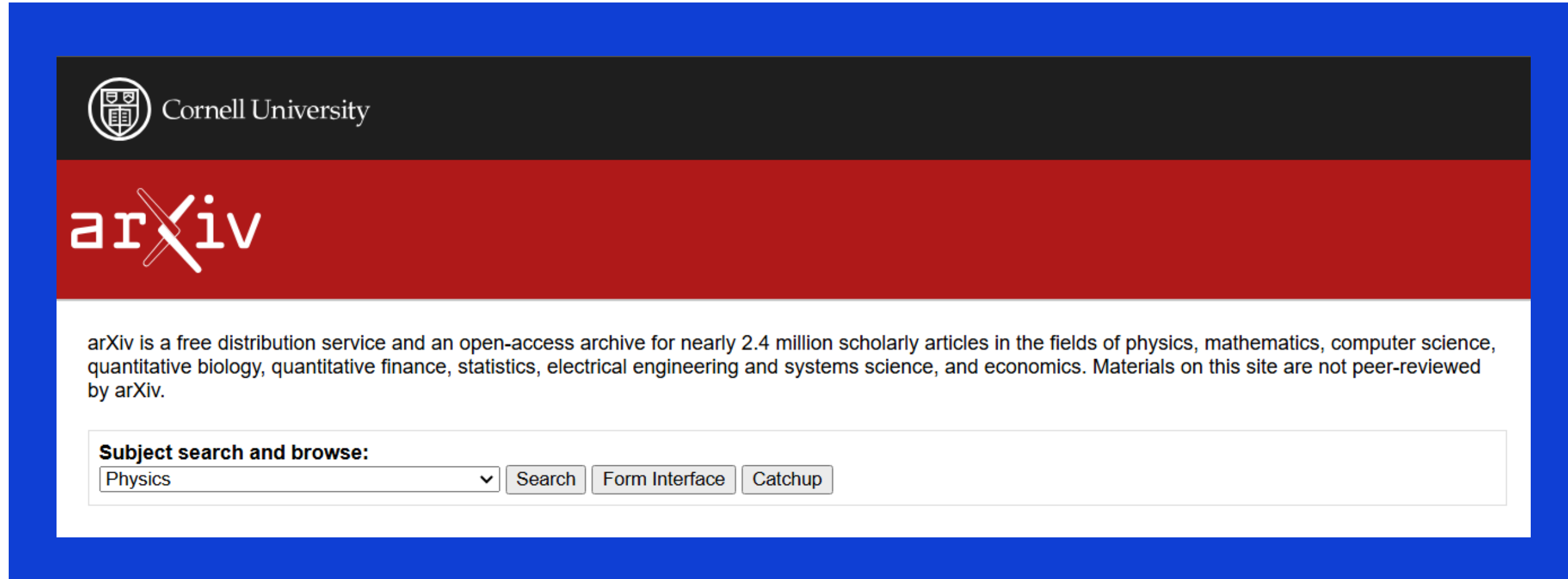
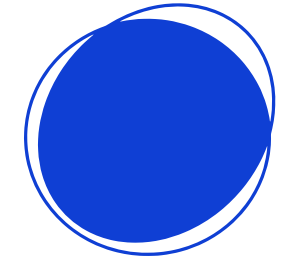
- Stworzenie narzędzia do generowania tytułów
- Projekt oparty na repozytorium `gpt-paper-title-generator`
- Przetestowanie rozwiązania

Przygotowanie projektu



- gpt-paper-title-generator
- Stworzenie zbioru danych
- Pobranie wstępnie wytrenowanego modelu AI
- Finetuning modelu
- Testy

Baza danych arXiv



- Scraping
- arxivscraper

GPT2

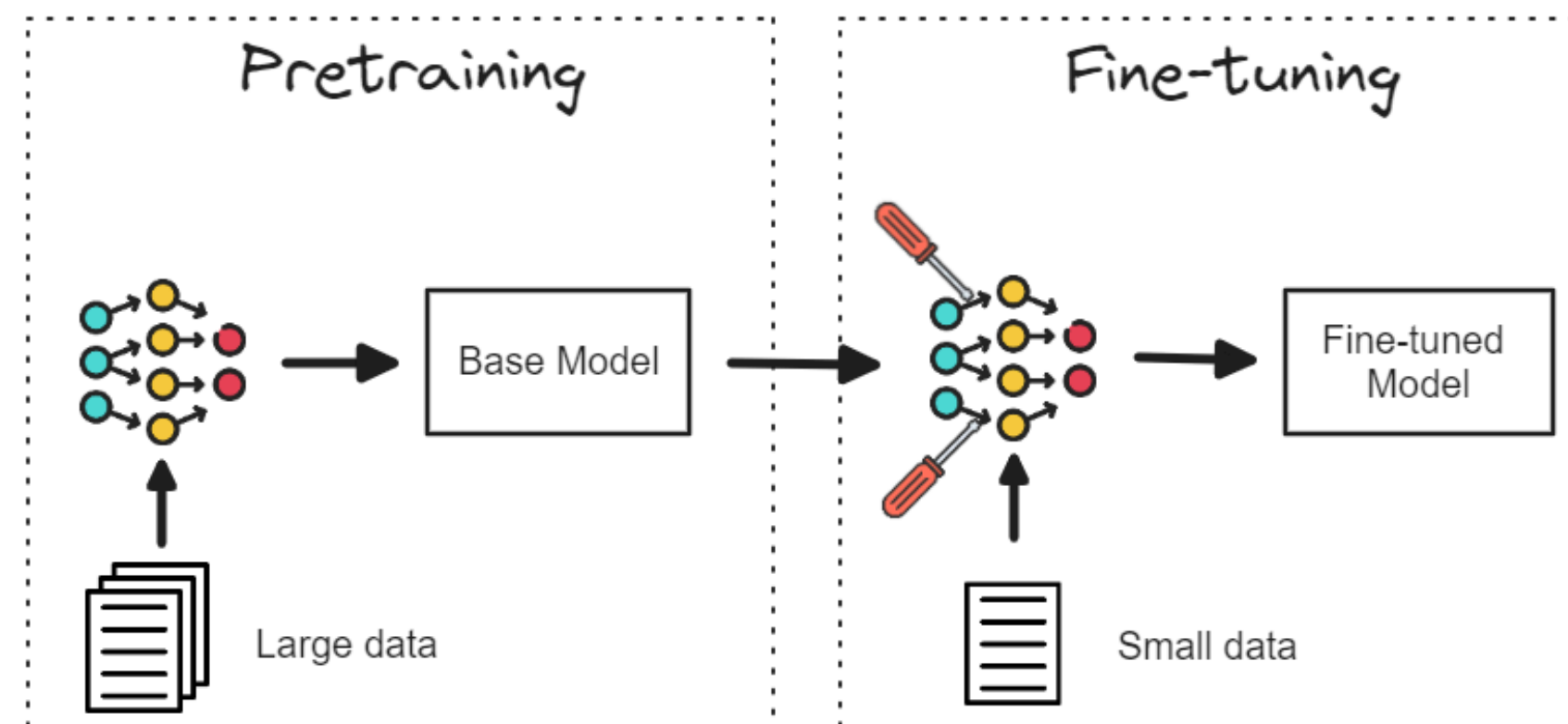


- **Generative Pre-training Transformer**
- **Transformer**
- **Model 124M**
- **Biblioteka gpt-2-simple**

Finetuning

- Dostrajanie wstępnie wytrenowanego modelu
- Mniejsze zasoby danych
- Training loss

Large Language Model



© Medium, Prasad Mahamulkar

Przygotowanie danych

```
scraper = ax.Scraper(category='eess', date_from='2024-05-01',  
                    date_until='2024-12-03', t=10,  
                    filters={'categories': ['eess.AS']})  
  
output = scraper.scrape()  
  
titles = [' '.join(o['title'].split()) for o in output]  
np.savetxt('titles_ref.csv', np.array(titles), fmt='%s')
```

Pobranie danych z kategorii „Electrical Engineering and Systems Science” i podkategorii „Audio and Speech Processing”, oraz sformatowanie danych.

Przykładowy plik .csv

dichotic harmony for the musical practice	✓ 1137
convolutional neural network achieves human-level accuracy in music genre classification	
acoustic scene classification: a competition review	
interfacing pdm <u>mems</u> microphones with pfm spiking systems: application for <u>neuromorphic</u> auditory sensors	
periodicity pitch detection in complex harmonies on eeg timeline data	
bayesian restoration of audio degraded by low-frequency pulses modeled via gaussian process	
sound field translation and mixed source model for virtual applications with perceptual validation	
version control of speaker recognition systems	
jointly fine-tuning "bert-like" self supervised models to improve multimodal speech emotion recognition	
ultra-low power on-chip learning of speech commands with phase-change memories	
<u>fullsubnet</u> : a full-band and sub-band fusion model for real-time single-channel speech enhancement	
data augmentation for end-to-end code-switching speech recognition	
enhancement by <u>postfiltering</u> for speech and audio coding in ad-hoc sensor networks	
bayesian learning for deep neural network adaptation	
adjust-free adversarial example generation in speech recognition using evolutionary multi-objective optimization under black-box condition	

Finetuning modelu

```
sess = gpt2.start_tf_sess()
gpt2.finetune(sess,
               'titles_ref.csv',
               model_name=model_name,
               steps=500,
               save_every=50,
               sample_every=25)

gpt2.generate(sess)
```

```
[101 | 2493.80] loss=1.38 avg=1.83
[102 | 2519.27] loss=1.58 avg=1.82
[103 | 2543.23] loss=1.45 avg=1.82
[104 | 2567.52] loss=1.56 avg=1.81
[105 | 2591.52] loss=1.37 avg=1.81
[106 | 2615.34] loss=1.48 avg=1.80
[107 | 2639.35] loss=1.48 avg=1.80
[108 | 2663.23] loss=1.22 avg=1.79
[109 | 2686.89] loss=1.41 avg=1.78
[110 | 2711.27] loss=1.50 avg=1.78
[111 | 2735.15] loss=1.37 avg=1.77
[112 | 2758.99] loss=1.30 avg=1.76
[113 | 2782.71] loss=1.40 avg=1.76
[114 | 2806.57] loss=1.30 avg=1.75
[115 | 2830.63] loss=1.26 avg=1.75
[116 | 2854.93] loss=1.31 avg=1.74
[117 | 2879.59] loss=1.35 avg=1.73
[118 | 2904.36] loss=1.33 avg=1.73
[119 | 2928.66] loss=1.44 avg=1.72
[120 | 2953.23] loss=1.28 avg=1.72
[121 | 2977.82] loss=1.18 avg=1.71
[122 | 3002.30] loss=1.31 avg=1.70
[123 | 3026.70] loss=1.10 avg=1.70
[124 | 3051.02] loss=1.47 avg=1.69
[125 | 3075.48] loss=1.13 avg=1.68
```

“Sample”

```
===== SAMPLE 1 =====
```

```
text>
```

```
<|startoftext>mamba: state-space modeling for multi-level sound classification<|endoftext|>
```

```
<|startoftext>the voice2i of chinese deepfake audio dataset<|endoftext|>
```

```
<|startoftext>an acoustic baseline for sound event localization via latent diffusion<|endoftext|>
```

```
<|startoftext>deepfake speech detector pooling<|endoftext|>
```

```
<|startoftext>japanese deepfake audio dataset: measuring the computational power of speaker recognition with low-rank adaptation and noisy diffusion<|endoftext|>
```

```
<|startoftext>multilingual speaker diarization for automatic speakers from different languages with cross-speaker retrieval<|endoftext|>
```

```
<|startoftext>self-supervised speech representation learning on speaker diarization and attention<|endoftext|>
```

```
<|startoftext>exploring a language model for the multilingual asr challenge 2022<|endoftext|>
```

```
<|startoftext>soundxml: introducing asr with style transform<|endoftext|>
```

```
<|startoftext>improving high-quality asr in monaural meetings: lessons from real-world meetings<|endoftext|>
```

```
<|startoftext>exploration of an asr-based language model for tts systems<|endoftext|>
```

```
<|startoftext>self-supervised transcription for neural end-to-end fast speech recognition<|endoftext|>
```

```
<|startoftext>enhancing speech analysis performance with text selective tts: real-world efficiency tuning<|endoftext|>
```

Generowanie tytułów

```
import gpt_2_simple as gpt2
sess = gpt2.start_tf_sess()
gpt2.load_gpt2(sess)

text = gpt2.generate(sess,
                      length=40,
                      temperature=0.5,
                      prefix=None,
                      nsamples=10,
                      batch_size=1,
                      return_as_list=True
                      )

for t in text:
    print("title ---")
    chk = t.title()
    chk = chk.replace('<|Startoftext|>', '').replace('\n', '') # remove extraneous stuff

    chk = chk[:chk.index('<|Endoftext|>')]
    print(chk)
    print("----")
```

Przykładowe tytuły bez prefiksu

```
title ---  
Self-Supervised Speech Representations Are More Interpretable  
---  
title ---  
LibriTTS-P: A Corpus With Speaking Style And Speaker Identity Prompts For Text-To-Speech And Style Captioning  
---  
title ---  
>Multimodal Punctuation By Speaker Characteristic Measure  
---  
title ---  
Leveraging Self-Supervised Models For Automatic Whispered Speech Recognition  
---  
title ---  
>Diffusion Gaussian Mixture Audio Denoise  
---  
title ---  
>Dynamic Humtrans: Humming Transcription Using Cnns And Dynamic Programming  
---
```

Prefiks “Neural”

```
title ---  
Neural Speech Enhancement In Competitively-Spaced Multi-Channel Multi-Speaker Environment Using Discrete Units  
---  
title ---  
Neuralized Spatial Learning For Audio Spoof Detection  
---  
title ---  
Neural Cues In Speech Signal Processing: A Novel Data Augmentation Approach  
---  
title ---  
Neuralized Text-To-Speech Synthesis With Controllable Semantic Representations  
---  
title ---  
Neural Signature Detection For Speaker Verification  
---  
title ---  
Neural Signature Generation Using Large Language Models  
---
```

Prefiks “Music”

```
title ---
Music|>Soundsil-Ds: Deep Denoising And Denoising Ofsea Change Usingilses For Audio Classification
---
title ---
Music|>Musebarrelmusic: An Age-Restricted Dataset For Chord-Based Music Generation
---
title ---
Music For All Platforms
---
title ---
Music Speech Representation Learning
---
title ---
Music|>Gibberish Is All You Need For Membership Verification In Fb
---
title ---
Music|>Sam1: Speaker Adaptive Mixture Of Experts For End-To-End Asr
---
title ---
Music>Mixture Of Experts Fusion For Llm-Based Asr With Enhanced Mems Microphone Arrays And Mems-Conformers
---
title ---
Music
---
```

Podsumowanie

- Projekt działa i generuję faktyczne tytuły
- Generowane tytuły nie są idealne, ale zdarzają się dobre przykłady
- Narzędzie mogłoby być bardziej użyteczne przy większym zbiorze danych i lepszym dostrojeniu
- Dodanie odpowiedniego prefiksu może polepszyć efekt

Dziękuję za uwagę

Bibliografia

- Gpt-paper-title-generator - <https://github.com/csinva/gpt-paper-title-generator>
- arXiv - <https://arxiv.org/>
- gpt-2-simple - <https://github.com/minimaxir/gpt-2-simple>
- arxivscraper - <https://github.com/Mahdisadjadi/arxivscraper>
- What is data scraping and how can you use it? -
- <https://targetinternet.com/resources/what-is-data-scraping-and-how-can-you-use-it>
- Fine-Tuning the Model: What, Why, and How -
- <https://medium.com/@amanatulla1606/fine-tuning-the-model-what-why-and-how-e7fa52bc8ddf>
- <https://openai.com/index/better-language-models/>