Search queries (27.10.2024): Data source <https://pubmed.ncbi.nlm.nih.gov/>

**Step 1:** issue queries separately for the respective technology

**1.** (((virology) OR (epidemiology)) AND (("neural network") OR ("artificial neural network") OR ("machine learning model") OR ("feedforward neural network") OR ("neural net algorithm") OR ("multilayer perceptron") OR ("convolutional neural network") OR ("recurrent neural network") OR ("long short-term memory network") OR ("CNN") OR ("GRNN") OR ("RNN") OR ("LSTM")))

3,791 results | virology-neural-networks.csv with metadata | virology-neural-networks.csv with abstracts

**2.** (((virology) OR (epidemiology)) AND (("deep learning") OR ("deep neural networks")))

2,664 results | virology-deep-learning.csv with metadata | virology-deep-learning.csv with abstracts

**3.** (((virology) OR (epidemiology)) AND (("computer vision") OR ("vision model") OR ("image processing") OR ("vision algorithms") OR ("computer graphics and vision") OR ("object recognition") OR ("scene understanding")))

3,647 results | virology-computer-vision.csv with metadata | virology-computer-vision.csv with abstracts

**4.** (((virology) OR (epidemiology)) AND (("natural language processing") OR ("text mining") OR (NLP) OR ("computational linguistics") OR ("language processing") OR ("text analytics") OR ("textual data analysis") OR ("text data analysis") OR ("text analysis") OR ("speech and language technology") OR ("language modeling") OR ("computational semantics")))

2,237 results | virology-natural-language-processing.csv with metadata | virology-natural-language-processing.csv with abstracts

**5.** (((virology) OR (epidemiology)) AND (("generative artificial intelligence") OR ("generative AI") OR ("generative deep learning") OR ("generative models")))

75 results | virology-generative-AI.csv with metadata | virology-generative-AI.csv with abstracts

**6.** (((virology) OR (epidemiology)) AND (("transformer models") OR ("self-attention models") OR ("transformer architecture") OR (transformer) OR ("attention-based neural networks") OR ("transformer networks") OR ("sequence-to-sequence models")))

262 results | virology-transformer.csv with metadata | virology-transformer.csv with abstracts

**7.** (((virology) OR (epidemiology)) AND (("large language model") OR (llm) OR ("transformer-based model") OR ("pretrained language model") OR ("generative language model") OR ("foundation model") OR ("state-of-the-art language model")))

183 results | virology-llm.csv with metadata | virology-llm.csv with abstracts

**8.** (((virology) OR (epidemiology)) AND (("multimodal model") OR ("multimodal neural network") OR ("vision transformer") OR ("diffusion model") OR ("generative diffusion model") OR ("diffusion-based generative model") OR ("continuous diffusion model")))

121 results | virology-vision-transformer.csv with metadata | virology-vision-transformer.csv with abstracts

—----------------------------------------------------------------------------------------------------------------------

There’s 12,980 total records with duplicates

**Step 2:** merge results and de-duplicate

There’s 11,450 records after aggregating and deduplication based on PMID

**Step 3:** fetch abstracts

result: [collection\_with\_abstracts.csv](https://github.com/jd-coderepos/virology-ai-papers/blob/main/collection_with_abstracts.csv)