

APPENDIX

Search terms

PubMed (n=125, 5 abstracts screened, 1 potential inclusion)

(((((("Venous Thromboembolism"[Mesh] OR "Pulmonary Embolism"[Mesh] OR "Venous Thrombosis"[Mesh])

OR ("venous thromboembolism"[tiab] OR "pulmonary embolism"[tiab] OR "deep vein thrombosis"[tiab] OR VTE[tiab] OR PE[tiab] OR DVT[tiab]))

)

AND

((("Neoplasms"[Mesh] OR "Cancer"[tiab] OR "Malignancy"[tiab] OR "Neoplasm"[tiab] OR "Oncologic"[tiab] OR "Cancer-Associated Thrombosis"[tiab] OR CAT[tiab]))

)

AND

((("Natural Language Processing"[Mesh] OR "Machine Learning"[Mesh] OR "Artificial Intelligence"[Mesh] OR "Text Mining"[Mesh]

OR "natural language processing"[tiab] OR "text mining"[tiab] OR NLP[tiab] OR "rule-based"[tiab] OR "machine learning"[tiab]

OR "deep learning"[tiab] OR "neural network"[tiab] OR "transformer"[tiab] OR "BERT"[tiab] OR "large language model*"[tiab]

OR "language model*"[tiab] OR LLM[tiab]))

)

)

NOT (Editorial[pt] OR Letter[pt])

AND (humans[MeSH Terms])

AND (english[lang])

EMBASE (n=380, 9 abstracts screened, 1 definite inclusion, 2 potential inclusions)

('venous thromboembolism'/exp OR 'pulmonary embolism'/exp OR 'deep vein thrombosis'/exp OR 'venous thromboembolism':ab,ti OR 'pulmonary embolism':ab,ti OR 'deep vein thrombosis':ab,ti OR VTE:ab,ti OR PE:ab,ti OR DVT:ab,ti)

AND

('neoplasm'/exp OR 'cancer'/exp OR 'malignant tumor'/exp OR 'cancer patient'/exp

OR 'cancer associated thrombosis'/exp OR cancer:ab,ti OR malignancy:ab,ti OR neoplasm:ab,ti OR 'oncologic patient*':ab,ti OR CAT:ab,ti)

AND

('natural language processing'/exp OR 'text mining'/exp OR 'machine learning'/exp

OR 'deep learning'/exp OR 'artificial intelligence'/exp OR 'transformer model'/exp

OR 'bert model'/exp OR 'large language model'/exp

OR 'natural language processing':ab,ti OR NLP:ab,ti OR 'text mining':ab,ti OR 'rule-based':ab,ti

OR 'machine learning':ab,ti OR 'deep learning':ab,ti OR 'transformer':ab,ti OR 'bert':ab,ti

OR 'large language model*':ab,ti OR LLM:ab,ti OR 'masked language model*':ab,ti)

NOT ([conference abstract]/lim OR [review]/lim OR [editorial]/lim OR [letter]/lim)

AND [english]/lim

AND [humans]/lim

AND [2010-2025]/py

Web of Science (n=284, 14 abstracts screened, 4 definite inclusions, 5 potential inclusions)

TS=("venous thromboembolism" OR "pulmonary embolism" OR "deep vein thrombosis" OR VTE OR PE OR DVT)

AND

(cancer OR malignancy OR neoplasm OR "oncologic patient*" OR "cancer-associated thrombosis" OR CAT)

AND

("natural language processing" OR NLP OR "text mining" OR "rule-based" OR "machine learning" OR "deep learning"

OR "neural network" OR "transformer" OR "BERT" OR "large language model*" OR "masked language model*" OR LLM))

Document types: Article

Language: English

Timespan: 2010–2025

Research domains: Medical Informatics, Oncology, Computer Science, Radiology