# Training

## Author Records

Pair of author records
In block k

Record # 1

Record # 2

name, title, abstract, label

Over 2.5 million pairs

## Representation

Author name

Abstract and title

Chars2vec

Character-level encoder

SPECTER

Semantic encoder

## NAND Model

Embedding #1

Encoded Record #1

Feed Forward NN

Shared Weights

Cosine Similarity

Encoded Record #2

Feed Forward NN

Embedding #2

# Inference

Block of author records
(e.g., J. Smith)

NAND Encoder

DBSCAN with constraints