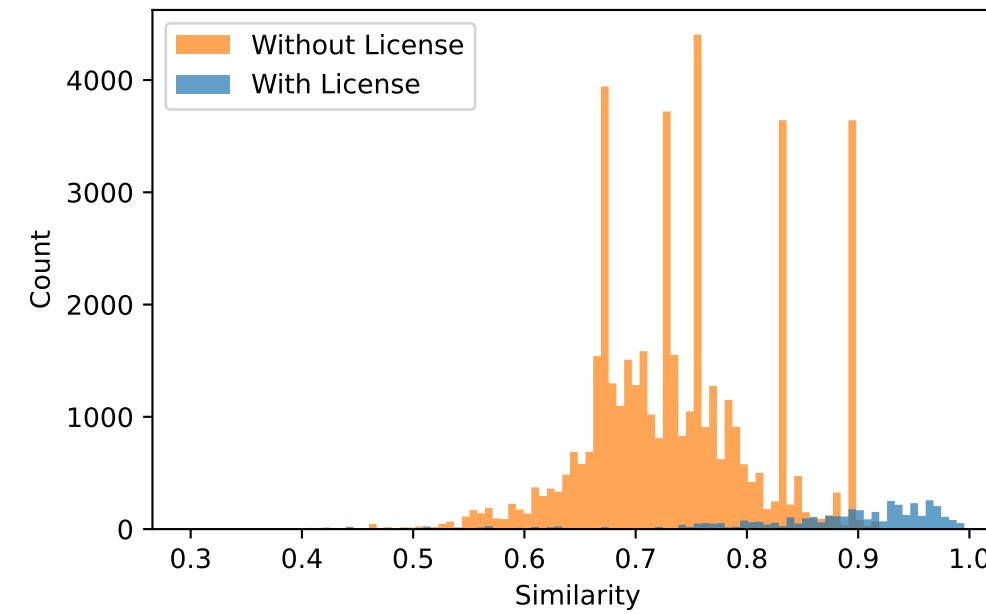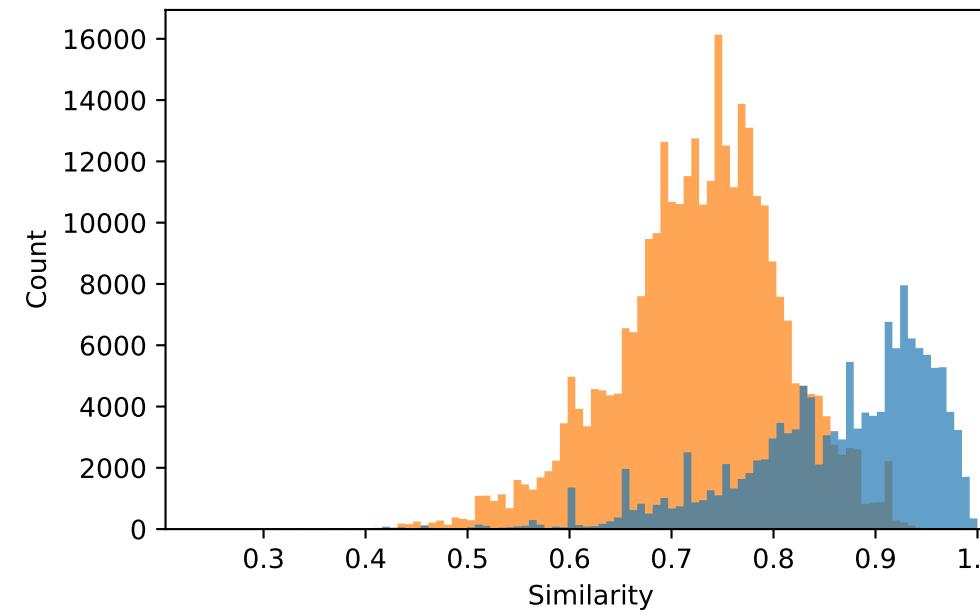Embedding Similarity Distributions by Article Quality Class and License Status