Key2Vec: Automatic Ranked Keyphrase Extraction from Scientific Articles using Phrase Embeddings - Supplementary Material

1 Phrase Embedding Sample Results

Table 1. shows some sample results obtained using the *Fasttext* model that we trained using the following parameters:

- learning rate = 0.05
- change the rate of updates for the learning rate = 100
- size of phrase vectors = 100
- size of the context window = 5
- number of epochs = 10
- minimal number of word occurences = 1
- number of negatives samples = 5
- max length of word ngram = 1
- loss function ns, hs, softmax = ns (negative sampling)
- number of buckets = 2000000
- min length of char ngram = 3
- max length of char ngram = 6
- number of threads = cpu_count of the 8 core machine
- sampling threshold = 0.0001

2 Corpus Statistics

Table 2. shows some statistics for the two benchmark datasets (*Inspec* and *SemEval 2010*) that we used for evaluation.

Corpus Statistic	**Inspec**	**SemEval**
Type	Abstracts	Full Articles
No. of Documents	500	100
Avg No. of Unigram Tokens	136.3	5179.6
Total No. of Annotated Keyphrases	4913	3003
Avg No. of Annotated Keyphrases	9.82	30.03
Total No. of Candidates for Key2Vec	6100	47159
Avg No. of Candidates for Key2Vec	12.2	471.59
Total No. of Matches	3562	958
Total Accuracy	72.50%	31.90%

Table 2: Corpus Statistics for the Benchmark datasets.

3 Sample Input and Output

Tables 3 - 7 shows samples from the *Inspec* dataset. It shows the input given to *Key2Vec* in the form of text content comprising of the whole text of a scientific article, and the output obtained by applying the *Key2Vec* procedure on the input. We don't show samples from SemEval 2010 dataset as they contain multi-page articles.

Phrase	Top 5 Similar Phrases
convolutional_neural_network	cnn, feature_representations, deep_convolutional_neural_network, deep_neural_network, scene_recognition
dark_matter	dm, dark_matter_particle, non-baryonic_dark_matter, dark_energy, self-interacting_dark_matter
natural_language_processing	nlp, language_processing, machine_translation, named_entity_recognition, sense_disambiguation
rnn	blstm, long_short-term_memory, lstms, handwritten_documents, recurrent_neural_network, lstm
svm	support_vector_machine, support_vector_machines, random_forest, svms, naive_bayes

Table 1: Top 5 similar phrases to a given phrase as produced by the phrase embedding model.

Title	**Compatibility of systems of linear constraints over the set of natural numbers**
Abstract	Criteria of compatibility of a system of linear Diophantine equations, strict inequations and nonstrict inequations are considered. Upper bounds for components of a minimal set of solutions and algorithms of construction,of minimal generating sets of solutions for all types of systems are given. These criteria and the corresponding algorithms for constructing a minimal supporting set of solutions can be used in solving all the considered types of systems and systems of mixed types.
Annotated Keyphrases	linear constraints, set of natural numbers, linear diophantine equations, strict inequations, nonstrict inequations, upper bound, minimal generating sets
Key2Vec Ranked Keyphrases	*linear constraints, natural numbers, strict,inequations, nonstrict inequations, upper bounds, linear diophantine equations, minimal set, corresponding algorithms, minimal generating sets, minimal supporting set, mixed types*

Table 3: Sample 1.

Title	**Identification of states of complex systems with estimation of admissible measurement errors on the basis of fuzzy information**
Abstract	The problem of identification of states of complex systems on the basis of, fuzzy values of informative attributes is considered. Some estimates of, a maximally admissible degree of measurement error are obtained that make it possible, using the apparatus of fuzzy set theory, to correctly,identify the current state of a system.
Annotated Keyphrases	complex system state identification, admissable measurement errors, fuzzy information, informative attribute, measurement error, fuzzy set theory
Key2Vec Ranked Keyphrases	*complex systems, fuzzy information, admissible,measurement errors, fuzzy values, informative attributes, measurement error, maximally admissible degree, fuzzy set theory, current state*

Table 4: Sample 2.

Title	**Nonlinear extrapolation algorithm for realization of a scalar random process**
Abstract	A method of construction of a nonlinear extrapolation algorithm is proposed. This method makes it possible to take into account any nonlinear random dependences that exist in an investigated process and are described by mixed central moment functions. The method is based on the V. S.,Pugachev canonical decomposition apparatus. As an example, the problem of nonlinear extrapolation is solved for a moment function of third order.
Annotated Keyphrases	nonlinear extrapolation algorithm, scalar random process, nonlinear random dependences, mixed central moment functions, canonical decomposition apparatus, moment functions
Key2Vec Ranked Keyphrases	*nonlinear extrapolation algorithm, scalar random process, investigated process, nonlinear random dependences, nonlinear extrapolation, moment function, mixed central moment functions*

Table 5: Sample 3.

Title	**Precoded OFDM with adaptive vector channel allocation for scalable video transmission over frequency-selective fading channels**
Abstract	Orthogonal frequency division multiplexing (OFDM) has been applied in broadband wireline and wireless systems for high data rate transmission where severe intersymbol interference (ISI) always occurs. The conventional,OFDM system provides advantages through conversion of an ISI channel into ISI-free subchannels at multiple frequency bands. However, it may suffer from channel spectral nulls and heavy data rate overhead due to cyclic prefix insertion. Previously, a new OFDM framework, the precoded,OFDM, has been proposed to mitigate the above two problems through precoding and conversion of an ISI channel into ISI-free vector channels. In this paper, we consider the application of the precoded,OFDM system to efficient scalable video transmission. We propose to,enhance the precoded OFDM system with adaptive vector channel allocation to provide stronger protection against errors to more,important layers in the layered bit stream structure of scalable video. The more critical layers, or equivalently, the lower layers, are,allocated vector channels of higher transmission quality. The channel quality is characterized by Frobenius norm metrics; based on channel estimation at the receiver. The channel allocation information is fed back periodically to the transmitter through a control channel. Simulation results have demonstrated the robustness of the proposed scheme to noise and fading inherent in wireless channels.
Annotated Keyphrases	precoded ofdm, scalable video transmission, frequency-selective fading channels, orthogonal frequency division multiplexing, channel spectral nulls, heavy data rate overhead, isi channel, isi-free vector channels, adaptive vector channel allocation, layered bit stream structure, lower layers, critical layer, channel quality, frobenius norm metrics, channel estimation, channel allocation information, control channel, robust
Key2Vec Ranked Keyphrases	*precoded ofdm, ofdm, frequency-selective fading channels, isi-free subchannels, isi channel, severe intersymbol interference, adaptive vector channel allocation, multiple frequency bands, channel spectral nulls, wireline and wireless systems, cyclic prefix insertion, isi-free vector channels, heavy data rate overhead, lower layers, vector channels, channel quality, simulation results, efficient scalable video transmission, proposed scheme, higher transmission quality, wireless channels, adaptive vector channel allocation, frobenius norm metrics, channel allocation information, layered bit stream structure, control channel*

Table 6: Sample 4.

Title	**Descriptological foundations of programming**
Abstract	Descriptological foundations of programming are constructed. An explication of the concept of a descriptive process is given. The operations of introduction and elimination of abstraction at the level of processes are refined. An intensional concept of a bipolar function is introduced. An explication of the concept of introduction and extraction of abstraction at the bipole level is given. On this basis a complete set of descriptological operations is constructed.
Annotated Keyphrases	descriptological foundations, programming, descriptive process, intensional concept, bipolar function, bipole level
Key2Vec Ranked Keyphrases	*descriptological foundations, programming descriptological foundations, descriptive process, intensional concept, bipolar function, bipole level, complete set, descriptological operations*

Table 7: Sample 5.