

Пример визуализации функции позиционного кодирования

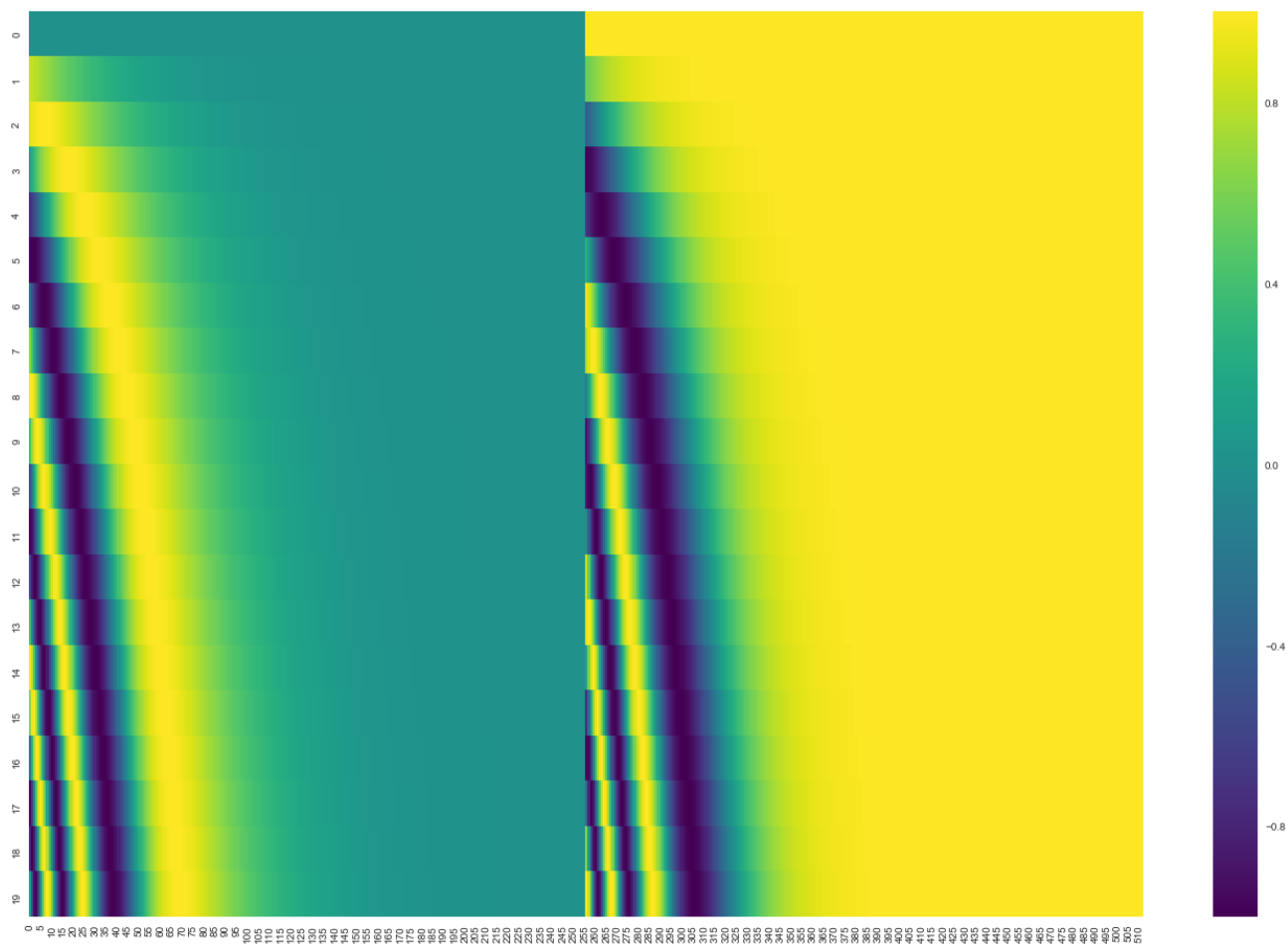


Рис. А1. Пример визуализации функции позиционного кодирования

Ссылки на исходные коды

Ссылка на github репозиторий https://github.com/andrsolo21/hse_Af_Tr_BERT.

Ссылка на ноутбук с примером работы модели BERT: [new_hse_BERT_0.ipynb](#).

Ссылка на ноутбук с примером работы модели ELMo: [new_hse_ELMO_0.ipynb](#).

Ссылка на ноутбук с агрегацией результатов: [Visualization.ipynb](#).

Визуализация эмбедингов

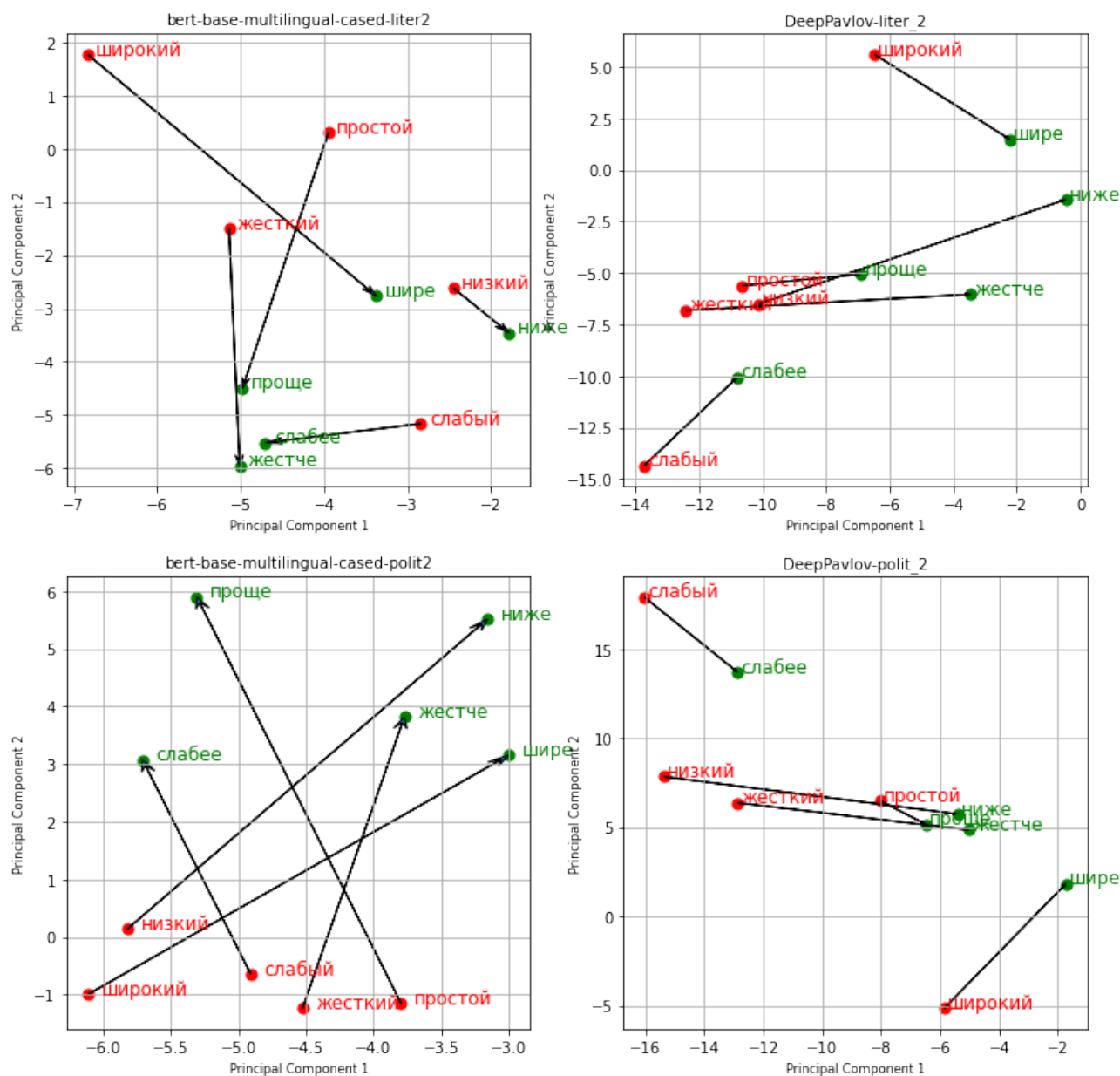


Рис. В1. Визуализация эмбедингов из группы «прилагательное – сравнительная степень» для модели BERT с обработкой N-грамм через вычисление среднего эмбединга; Слева модель bert-base-multilingual-cased, справа от DeepPavlov; Верхние модели обучены на жанре литературы, нижние на политике

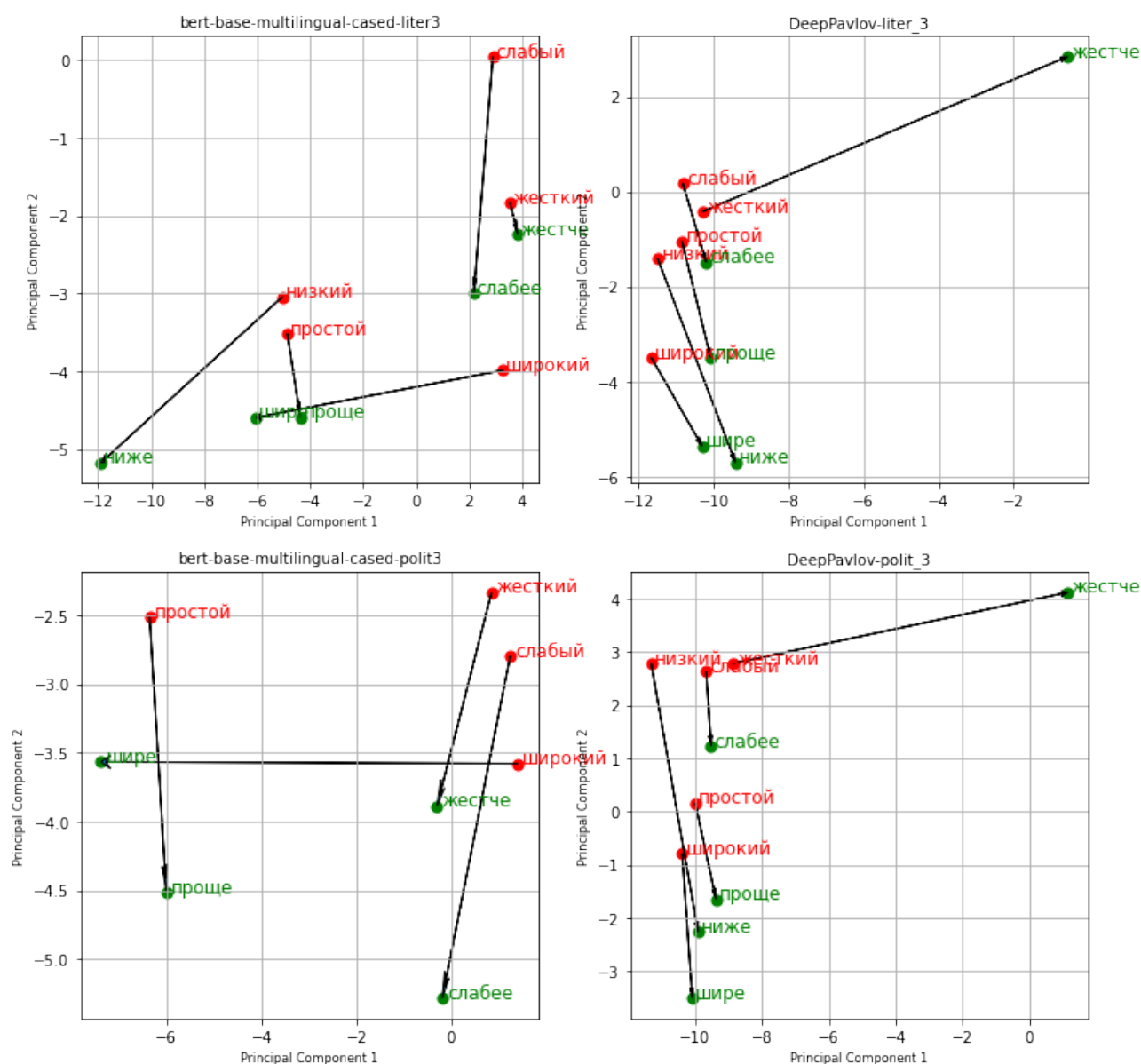


Рис. В2. Визуализация эмбедингов из группы «прилагательное – сравнительная степень» для модели BERT с обработкой N-грамм через вычисление суммы эмбедингов; Слева модель bert-base-multilingual-cased, справа от DeepPavlov; Верхние модели обучены на жанре литературы, нижние на политике

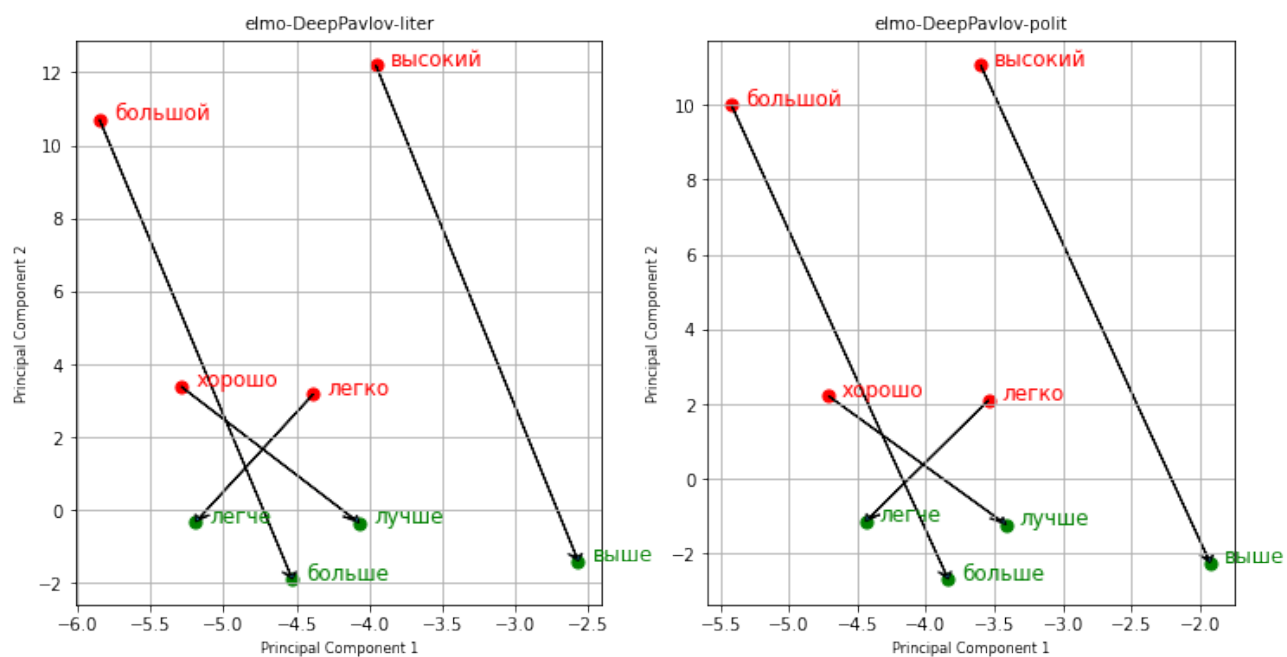


Рис. В3. Визуализация эмбедингов из группы «прилагательное – сравнительная степень» для модели ELMo; Слева модель обучена на жанре литературы, справа на политике

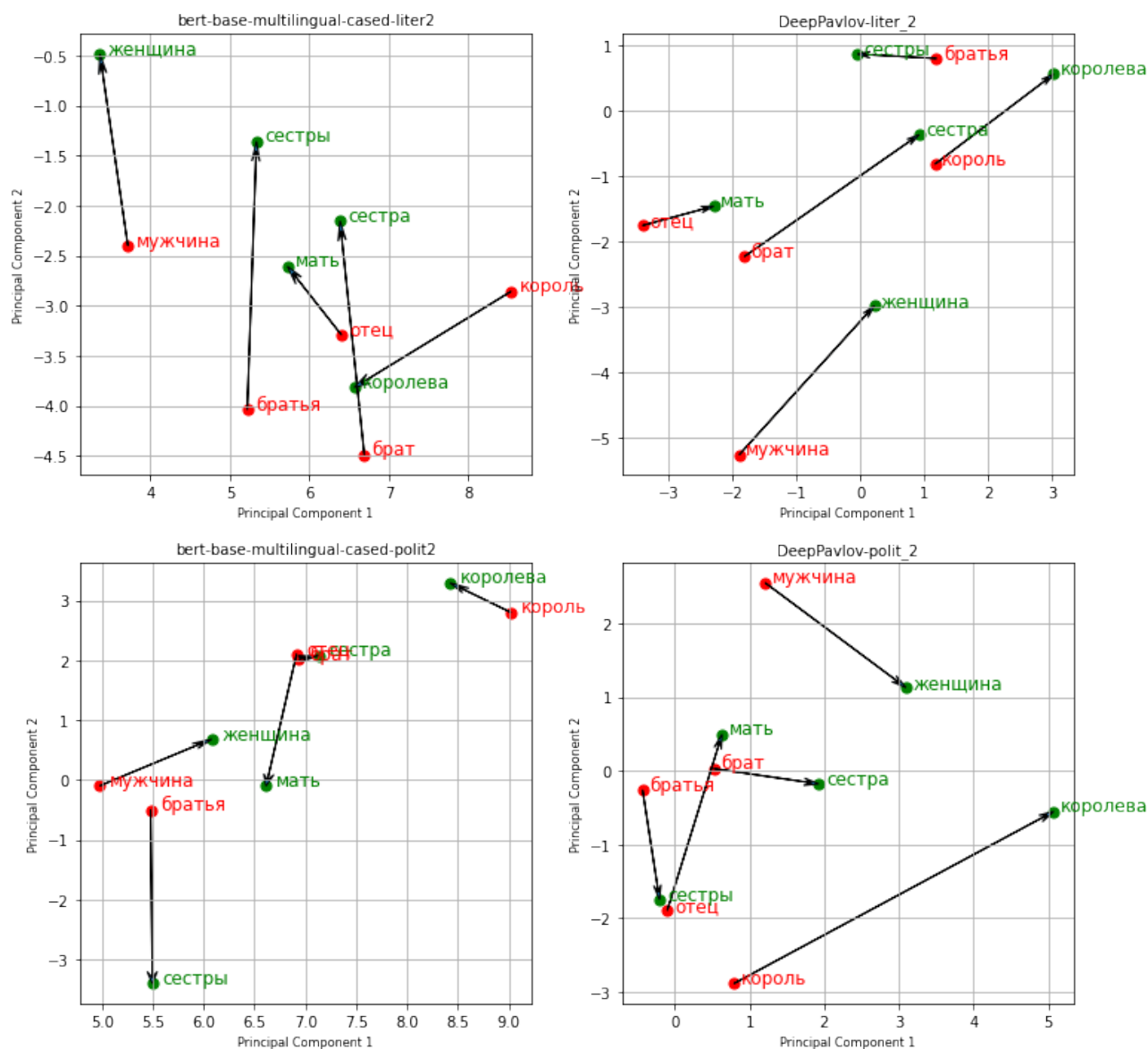


Рис. В4. Визуализация эмбедингов из группы «мужской пол – женский пол» для модели BERT с обработкой N-грамм через вычисление среднего эмбединга; Слева модель bert-base-multilingual-cased, справа от DeepPavlov; Верхние модели обучены на жанре литературы, нижние на политике

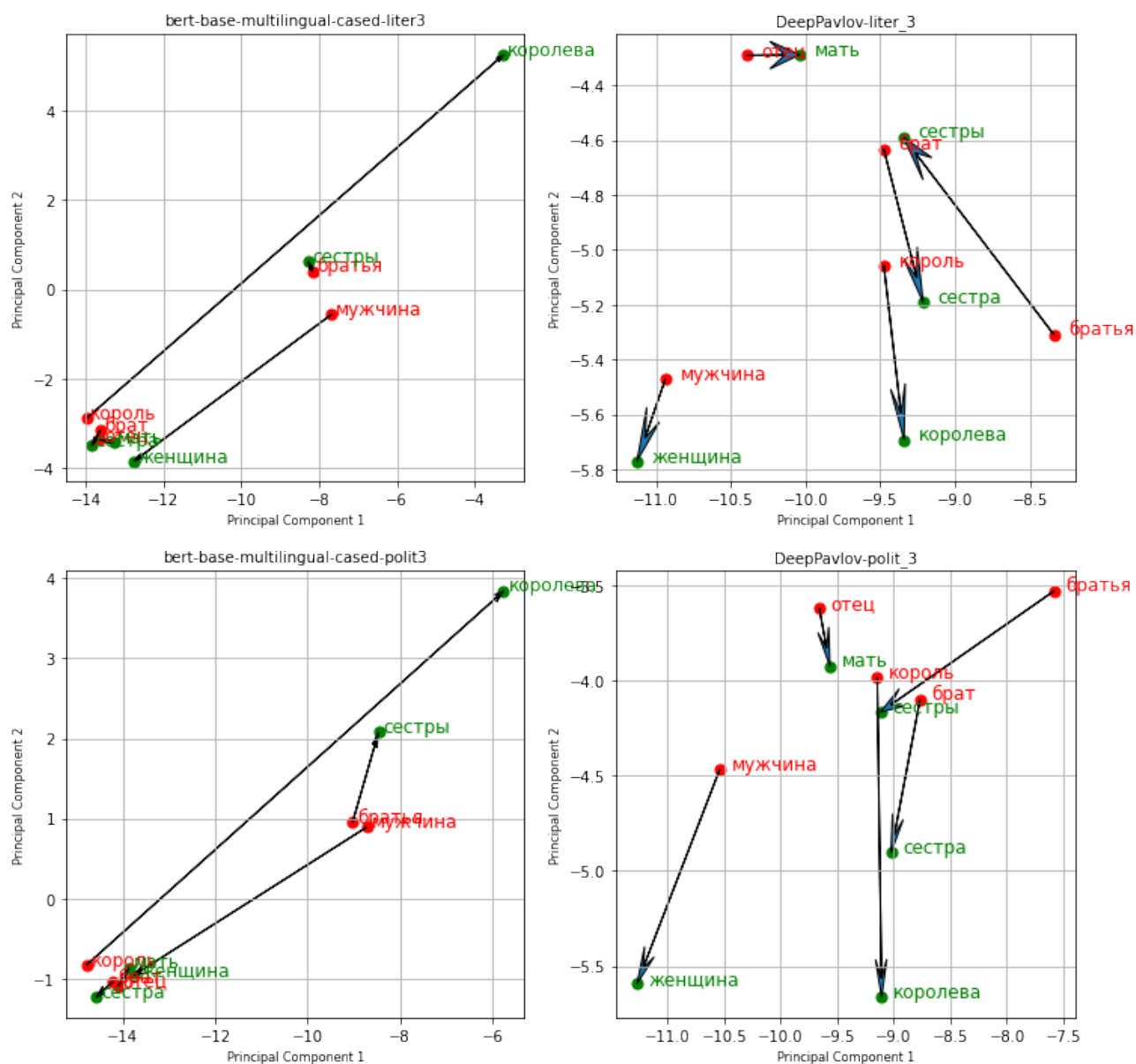


Рис. В5. Визуализация эмбедингов из группы «мужской пол – женский пол» для модели BERT с обработкой N-грамм через вычисление суммы эмбедингов; Слева модель bert-base-multilingual-cased, справа от DeepPavlov; Верхние модели обучены на жанре литературы, нижние на политике

Результаты тестирования моделей

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Мужской пол - женский пол	True	1.000000	True	0.900485	0.900485	0.942573	1.0	2
Прилагательное - сравнительная степень	False	0.333333	False	0.895160	0.941047	0.930116	1.0	6

Рис. Г1. Результаты тестирования для модели bert-base-multilingual-cased-liter

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.851062	0.887988	0.926910	1.000000	30
Мужской пол - женский пол	0.384615	0.416667	0.378205	0.883500	0.901546	0.950992	1.000000	156
Национальность - прилагательное	0.367470	0.433735	0.361446	0.873831	0.893227	0.932075	0.999993	332
Прилагательное - наречие	0.306324	0.395257	0.306324	0.877269	0.902574	0.960834	1.000000	506
Прилагательное - сравнительная степень	0.370370	0.436508	0.366402	0.877970	0.895814	0.958734	1.000000	756
Столицы - страны	0.000000	0.000000	0.000000	0.804135	0.847072	0.916737	0.996707	115

Рис. Г2. Результаты тестирования для модели bert-base-multilingual-cased-liter2

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.686737	0.711935	0.926910	1.000000	30
Мужской пол - женский пол	0.224359	0.288462	0.378205	0.786111	0.804744	0.950992	1.000000	156
Национальность - прилагательное	0.262048	0.340361	0.361446	0.804582	0.820299	0.932075	0.999992	332
Прилагательное - наречие	0.270751	0.359684	0.306324	0.906424	0.924559	0.960834	1.000000	506
Прилагательное - сравнительная степень	0.314815	0.392857	0.366402	0.845742	0.861733	0.958734	1.000000	756
Столицы - страны	0.000000	0.008696	0.000000	0.787862	0.832490	0.916737	0.996530	115

Рис. Г3. Результаты тестирования для модели bert-base-multilingual-cased-liter3

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Мужской пол - женский пол	0.5	0.500000	0.5	0.848026	0.878964	0.921125	1.0	2
Прилагательное - сравнительная степень	0.0	0.333333	0.0	0.882311	0.936110	0.919653	1.0	6

Рис. Г4. Результаты тестирования для модели bert-base-multilingual-cased-polit

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.033333	0.033333	0.033333	0.842050	0.879136	0.937145	1.000000	30
Мужской пол - женский пол	0.500000	0.550000	0.450000	0.865871	0.883225	0.945974	1.000000	20
Национальность - прилагательное	0.406299	0.505512	0.403150	0.884483	0.897690	0.938001	0.999998	635
Прилагательное - наречие	0.264069	0.367965	0.268398	0.866413	0.893137	0.950146	1.000000	462
Прилагательное - сравнительная степень	0.310000	0.383333	0.308333	0.866738	0.885623	0.944975	1.000000	600
Столицы - страны	0.001117	0.041341	0.001117	0.821504	0.860237	0.933539	0.999091	895

Рис. Г5. Результаты тестирования для модели bert-base-multilingual-cased-polit2

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.678521	0.711428	0.937296	1.000000	20
Мужской пол - женский пол	0.250000	0.250000	0.450000	0.674139	0.712448	0.945974	1.000000	20
Национальность - прилагательное	0.287179	0.374359	0.429060	0.823580	0.837703	0.938724	0.999997	585
Прилагательное - наречие	0.281579	0.402632	0.310526	0.900844	0.922219	0.948546	1.000000	380
Прилагательное - сравнительная степень	0.388889	0.479532	0.397661	0.854247	0.866328	0.945452	1.000000	342
Столицы - страны	0.000000	0.030093	0.001157	0.795631	0.833677	0.933368	0.998820	864

Рис. Г6. Результаты тестирования для модели bert-base-multilingual-cased-polit3

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.793130	0.819003	0.918517	1.000000	20
Мужской пол - женский пол	0.659091	0.757576	0.643939	0.917592	0.917783	0.937557	1.000000	132
Национальность - прилагательное	0.333333	0.333333	0.333333	0.899240	0.911733	0.968291	1.000000	6
Прилагательное - наречие	0.222222	0.460784	0.222222	0.857969	0.888816	0.930546	1.000000	306
Прилагательное - сравнительная степень	0.206522	0.351449	0.215580	0.845331	0.862445	0.907858	1.000000	552
Столицы - страны	0.666667	1.000000	0.666667	0.832437	0.893409	0.832437	0.921754	3

Рис. Г7. Результаты тестирования для модели DeepPavlov-liter

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.017857	0.000000	0.812737	0.837997	0.923055	1.000000	56
Мужской пол - женский пол	0.631868	0.730769	0.626374	0.915560	0.916735	0.940726	1.000000	182
Национальность - прилагательное	0.401961	0.500000	0.441176	0.860990	0.869891	0.974966	0.999993	408
Прилагательное - наречие	0.115942	0.259058	0.117754	0.862808	0.897032	0.944166	1.000000	552
Прилагательное - сравнительная степень	0.198276	0.305419	0.197044	0.844081	0.860322	0.910856	1.000000	812
Столицы - страны	0.029851	0.034826	0.044776	0.797999	0.842147	0.888389	0.998932	201

Рис. Г8. Результаты тестирования для модели DeepPavlov-liter_2

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.813344	0.834249	0.922889	1.000000	56
Мужской пол - женский пол	0.576923	0.686813	0.626374	0.838709	0.840142	0.940715	1.000000	182
Национальность - прилагательное	0.350490	0.397059	0.441176	0.905001	0.921356	0.974966	0.999995	408
Прилагательное - наречие	0.103261	0.230072	0.119565	0.874341	0.905416	0.944115	1.000000	552
Прилагательное - сравнительная степень	0.162698	0.265873	0.201058	0.758347	0.779599	0.910763	1.000000	756
Столицы - страны	0.038043	0.043478	0.043478	0.739623	0.781124	0.889273	0.999069	184

Рис. Г9. Результаты тестирования для модели DeepPavlov-liter_3

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.792315	0.812806	0.894300	1.000000	2
Мужской пол - женский пол	0.309524	0.380952	0.309524	0.888865	0.894410	0.923934	1.000000	42
Национальность - прилагательное	0.333333	0.500000	0.333333	0.906793	0.916743	0.975569	1.000000	12
Прилагательное - наречие	0.271930	0.467836	0.269006	0.856142	0.886683	0.929608	1.000000	342
Прилагательное - сравнительная степень	0.266082	0.321637	0.245614	0.841756	0.859379	0.896573	1.000000	342
Столицы - страны	0.000000	0.333333	0.000000	0.892919	1.000000	0.892919	0.928382	3

Рис. Г10. Результаты тестирования для модели DeepPavlov-polit

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.095238	0.142857	0.095238	0.805934	0.827822	0.925600	1.000000	42
Мужской пол - женский пол	0.500000	0.566667	0.466667	0.904818	0.909209	0.933786	1.000000	30
Национальность - прилагательное	0.404551	0.471555	0.457649	0.853775	0.862004	0.979647	0.999997	791
Прилагательное - наречие	0.145022	0.303030	0.151515	0.851967	0.888582	0.933824	1.000000	462
Прилагательное - сравнительная степень	0.184615	0.267692	0.184615	0.844823	0.859809	0.902036	1.000000	650
Столицы - страны	0.009730	0.051892	0.009730	0.828142	0.858813	0.934850	0.999316	925

Рис. Г11. Результаты тестирования для модели DeepPavlov-polit_2

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.035714	0.035714	0.000000	0.755076	0.769617	0.918404	1.000000	56
Мужской пол - женский пол	0.581818	0.681818	0.654545	0.855254	0.859082	0.938313	1.000000	110
Национальность - прилагательное	0.320819	0.412969	0.482935	0.909286	0.920953	0.980581	0.999998	586
Прилагательное - наречие	0.141538	0.267692	0.153846	0.870396	0.903275	0.943224	1.000000	650
Прилагательное - сравнительная степень	0.174603	0.253968	0.201058	0.762850	0.782350	0.905826	1.000000	756
Столицы - страны	0.004963	0.028536	0.008685	0.783193	0.811331	0.924831	0.998455	806

Рис. Г12. Результаты тестирования для модели DeepPavlov-polit_3

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.529524	0.711460	0.795986	1.000000	2
Национальность - прилагательное	1.000000	1.000000	1.000000	0.777929	0.777929	0.911426	0.997948	11
Прилагательное - наречие	0.800000	0.933333	0.800000	0.695183	0.710281	0.834578	1.000000	30
Прилагательное - сравнительная степень	0.214286	0.410714	0.214286	0.667024	0.728323	0.755446	1.000000	56

Рис. Г13. Результаты тестирования для модели elmo-DeepPavlov-liter

metric	accuracy			cosine				count
	3CosAdd	3CosAdd_3	3CosMul	3CosAdd	3CosAdd_3	3CosMul	3CosAvg	
Антонимы	0.000000	0.000000	0.000000	0.527943	0.711460	0.789720	1.0	2
Прилагательное - наречие	0.000000	0.500000	0.000000	0.739266	0.836530	0.773585	1.0	2
Прилагательное - сравнительная степень	0.214286	0.428571	0.214286	0.714386	0.760131	0.790758	1.0	42

Рис. Г14. Результаты тестирования для модели elmo-DeepPavlov-polit