

# HW1

xy990

October 2018

## 1 Hyper-parameters Tuning

### 1.1 Tokenize

For the parameter tokenizer tuning, I used 3 techniques, one is only remove punctuation, another is only remove stop words, another is stemming, besides, I also combine three methods together remove punctuation, stop words and stemming. Compared with validation accuracy, only the methods which only remove stop words performs best. Please see results in figure 1 and 2 .

### 1.2 N-grams

For the parameter N-grams tuning, I choose N-grams is 1, 2, 3, 4. Compared with the validation accuracy, unigram performs best. Please see results in figure 3 and 4 .

### 1.3 Optimizer

For the optimizer I choose to compare the performance of SGD and Adam, it is a little difficult to compare them because they have different performance with different magnitude of learning rate. Therefore, for SGD with momentum of 0.9, I choose the model with learning rate 0.01, 0.1 and 1. For Adam, I choose the model with learning rate 0.001, 0.01, 0.1. Due to early stopping, model with Adam and learning rate 0.01 early stops within 5 epochs, so I do not plot this result in the learning curve figures. For other choice, they do not meet early stopping, the results show Adam and learning rate 0.001 performs best, and

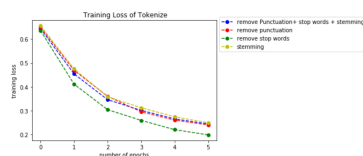


Figure 1: training loss of tokenizer

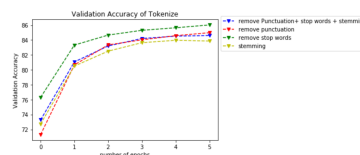


Figure 2: validation accuracy of tokenizer

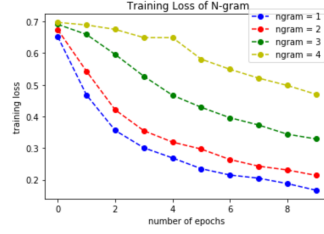


Figure 3: training loss of N-grams

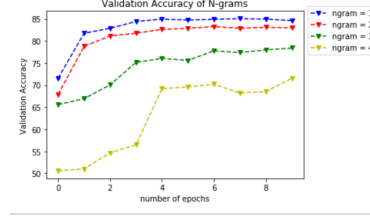


Figure 4: validation accuracy of N-grams

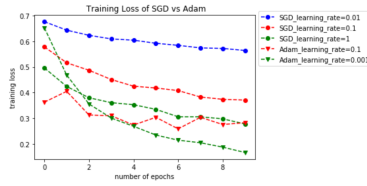


Figure 5: training loss of Optimizer(SGD vs Adam)

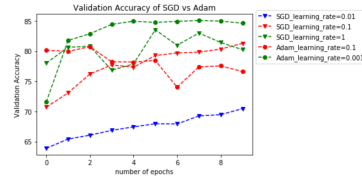


Figure 6: validation accuracy of Optimizer(SGD vs Adam)

SGD and learning rate 1 performs good, but it fluctuate in validation accuracy due to large learning rate. Please see results in figure 5 and 6 .

## 1.4 Learning Rate

For parameter learning rate, I choose SGD with momentum 0.9 as optimizer and learning rate 0.001, 0.01, 0.1, 1. Due to the very bad performance of learning rate 0.001 for SGD in 10 epochs, I do not plot and show it in the table. Compared with the validation accuracy, model with SGD and learning rate 1 performs best, but the validation accuracy fluctuates and does not converge. Please see results in figure 7 and 8 .

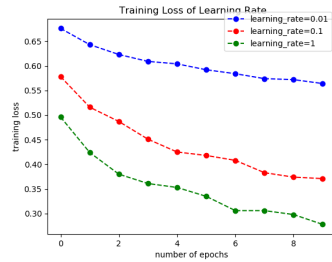


Figure 7: training loss of Learning Rate



Figure 8: validation accuracy of Learning Rate

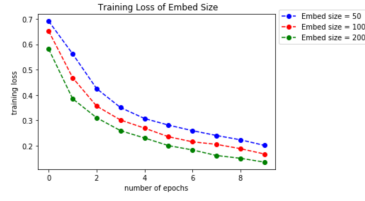


Figure 9: training loss of Embed Size

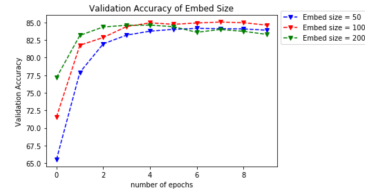


Figure 10: validation accuracy of Embed Size

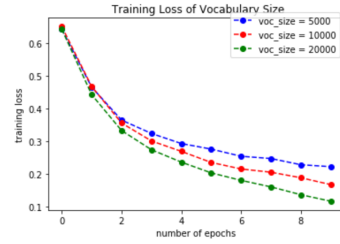


Figure 11: training loss of vocabulary size

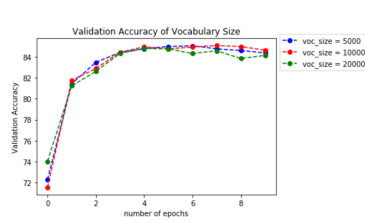


Figure 12: validation accuracy of vocabulary size

## 1.5 Embedding Size

For embedding size, I choose 50, 100 and 200. Compared with the validation accuracy, embedding size of 100 performs best. Please see results in figure 9 and 10 .

## 1.6 Vocabulary Size

For vocabulary size, I choose 5000, 10000 and 20000. I expect vocabulary size larger, the performance should be better. However, compared with the validation accuracy, vocabulary size of 5000 performs slightly better than others. Please see results in figure 11 and 12 .

## 1.7 Table of Training Loss and Validation Accuracy for All Parameters Tuning

For the results of all parameters tuning, please see table 13 and 14 .

## 2 Best Model and Error Analysis

After tuning the parameters, I choose remove stop words for tokenizer, 1 for N-grams, 5000 for vocabulary size, Adam for optimizer, 100 for embedding size and 0.001 for learning rate, the model reaches Val Acc 84.54 and Test Acc 85.8.

Template	Region	Model	Optimizer	method	loss	LearningRate	init	act1	act2	act3	act4	act5	act6	act7	act8	act9	act10	act11	act12	act13	act14	act15	act16	act17	act18	act19	act20	act21	act22	act23	act24	act25	act26	act27	act28	act29	act30	act31	act32	act33	act34	act35	act36	act37	act38	act39	act40	act41	act42	act43	act44	act45	act46	act47	act48	act49	act50	act51	act52	act53	act54	act55	act56	act57	act58	act59	act60	act61	act62	act63	act64	act65	act66	act67	act68	act69	act70	act71	act72	act73	act74	act75	act76	act77	act78	act79	act80	act81	act82	act83	act84	act85	act86	act87	act88	act89	act90	act91	act92	act93	act94	act95	act96	act97	act98	act99	act100	act101	act102	act103	act104	act105	act106	act107	act108	act109	act110	act111	act112	act113	act114	act115	act116	act117	act118	act119	act120	act121	act122	act123	act124	act125	act126	act127	act128	act129	act130	act131	act132	act133	act134	act135	act136	act137	act138	act139	act140	act141	act142	act143	act144	act145	act146	act147	act148	act149	act150	act151	act152	act153	act154	act155	act156	act157	act158	act159	act160	act161	act162	act163	act164	act165	act166	act167	act168	act169	act170	act171	act172	act173	act174	act175	act176	act177	act178	act179	act180	act181	act182	act183	act184	act185	act186	act187	act188	act189	act190	act191	act192	act193	act194	act195	act196	act197	act198	act199	act200	act201	act202	act203	act204	act205	act206	act207	act208	act209	act210	act211	act212	act213	act214	act215	act216	act217	act218	act219	act220	act221	act222	act223	act224	act225	act226	act227	act228	act229	act230	act231	act232	act233	act234	act235	act236	act237	act238	act239	act240	act241	act242	act243	act244	act245	act246	act247	act248	act249	act250	act251	act252	act253	act254	act255	act256	act257	act258	act259	act260	act261	act262	act263	act264	act265	act266	act267	act268	act269	act270	act271	act272	act273	act274	act275	act276	act277	act278	act279	act280	act281	act282	act283	act284	act285	act286	act287	act288	act289	act290	act291	act292	act293	act294	act295	act296	act297	act298	act299	act300	act301	act302	act303	act304	act305	act306	act307	act308	act309	act310	act311	act312	act313	act314	act315	act316	act317	act318	act319	act320	act321	act322	act323	act324	act325	act326	act327	act328	act329	act330	act331	act332	act333	act334	act335	act336	act337	act338	act339	act340	act341	act342	act343	act344	act345	act346	act347	act348	act349	act350	act351	act352	act353	act354	act355	act356	act357	act358	act359	act360	act361	act362	act363	act364	act365	act366	act367	act368	act369	act370	act371	act372	act373	act374	act375	act376	act377	act378	act379	act380	act381	act382	act383	act384	act385	act386	act387	act388	act389	act390	act391	act392	act393	act394	act395	act396	act397	act398	act399	act400	act401	act402	act403	act404	act405	act406	act407	act408	act409	act410	act411	act412	act413	act414	act415	act416	act417	act418	act419	act420	act421	act422	act423	act424	act425	act426	act427	act428	act429	act430	act431	act432	act433	act434	act435	act436	act437	act438	act439	act440	act441	act442	act443	act444	act445	act446	act447	act448	act449	act450	act451	act452	act453	act454	act455	act456	act457	act458	act459	act460	act461	act462	act463	act464	act465	act466	act467	act468	act469	act470	act471	act472	act473	act474	act475	act476	act477	act478	act479	act480	act481	act482	act483	act484	act485	act486	act487	act488	act489	act490	act491	act492	act493	act494	act495	act496	act497	act498	act499	act500	act501	act502	act503	act504	act505	act506	act507	act508	act509	act510	act511	act512	act513	act514	act515	act516	act517	act518	act519	act520	act521	act522	act523	act524	act525	act526	act527	act528	act529	act530	act531	act532	act533	act534	act535	act536	act537	act538	act539	act540	act541	act542	act543	act544	act545	act546	act547	act548	act549	act550	act551	act552	act553	act554	act555	act556	act557	act558	act559	act560	act561	act562	act563	act564	act565	act566	act567	act568	act569	act570	act571	act572	act573	act574	act575	act576	act577	act578	act579	act580	act581	act582	act583	act584	act585	act586	act587	act588	act589	act590	act591	act592	act593	act594	act595	act596	act597	act598	act599	act600	act601	act602	act603	act604	act605	act606	act607	act608	act609	act610	act611	act612	act613	act614	act615	act616	act617	act618	act619	act620	act621	act622	act623	act624	act625	act626	act627	act628	act629	act630	act631	act632	act633	act634	act635	act636	act637	act638	act639	act640	act641	act642	act643	act644	act645	act646	act647	act648	act649	act650	act651	act652	act653	act654	act655	act656	act657	act658	act659	act660	act661	act662	act663	act664	act665	act666	act667	act668	act669	act670	act671	act672	act673	act674	act675	act676	act677	act678	act679	act680	act681	act682	act683	act684	act685	act686	act687	act688	act689	act690	act691	act692	act693	act694	act695	act696	act697	act698	act699	act700	act701	act702	act703	act704	act705	act706	act707	act708	act709	act710	act711	act712	act713	act714	act715	act716	act717	act718	act719	act720	act721	act722	act723	act724	act725	act726	act727	act728	act729	act730	act731	act732	act733	act734	act735	act736	act737	act738	act739	act740	act741	act742	act743	act744	act745	act746	act747	act748	act749	act750	act751	act752	act753	act754	act755	act756	act757	act758	act759	act760	act761	act762	act763	act764	act765	act766	act767	act768	act769	act770	act771	act772	act773	act774	act775	act776	act777	act778	act779	act780	act781	act782	act783	act784	act785	act786	act787	act788	act789	act790	act791	act792	act793	act794	act795	act796	act797	act798	act799	act800	act801	act802	act803	act804	act805	act806	act807	act808	act809	act810	act811	act812	act813	act814	act815	act816	act817	act818	act819	act820	act821	act822	act823	act824	act825	act826	act827	act828	act829	act830	act831	act832	act833	act834	act835	act836	act837	act838	act839	act840	act841	act842	act843	act844	act845	act846	act847	act848	act849	act850	act851	act852	act853	act854	act855	act856	act857	act858	act859	act860	act861	act862	act863	act864	act865	act866	act867	act868	act869	act870	act871	act872	act873	act874	act875	act876	act877	act878	act879	act880	act881	act882	act883	act884	act885	act886	act887	act888	act889	act890	act891	act892	act893	act894	act895	act896	act897	act898	act899	act900	act901	act902	act903	act904	act905	act906	act907	act908	act909	act910	act911	act912	act913	act914	act915	act916	act917	act918	act919	act920	act921	act922	act923	act924	act925	act926	act927	act928	act929	act930	act931	act932	act933	act934	act935	act936	act937	act938	act939	act940	act941	act942	act943	act944	act945	act946	act947	act948	act949	act950	act951	act952	act953	act954	act955	act956	act957	act958	act959	act960	act961	act962	act963	act964	act965	act966	act967	act968	act969	act970	act971	act972	act973	act974	act975	act976	act977	act978	act979	act980	act981	act982	act983	act984	act985	act986	act987	act988	act989	act990	act991	act992	act993	act994	act995	act996	act997	act998	act999	act1000
----------	--------	-------	-----------	--------	------	--------------	------	------	------	------	------	------	------	------	------	------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	---------

Figure 14: validation accuracy for all parameters tuning

[illegible]

Figure 16: wrong reviews

correct predicted reviews and 3 wrong  
16 .

notebook and excel in Github. This is Phw1