

# SQuAD 2.0

## The Stanford Question Answering Dataset

### What is SQuAD?

---

Stanford **Question Answering Dataset** (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable.

**New** **SQuAD2.0** combines the 100,000 questions in SQuAD1.1 with over 50,000 new, unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering. SQuAD2.0 is a challenging natural language understanding task for existing models, and we release SQuAD2.0 to the community as the successor to SQuAD1.1. We are optimistic that this new dataset will encourage the development of reading comprehension systems that know what they don't know.

(/SQuAD-explorer/explore/v2.0/dev/)

(<http://arxiv.org/abs/1806.03822>)

**SQuAD 1.1**, the previous version of the SQuAD dataset, contains 100,000+ question-answer pairs on 500+ articles.

(/SQuAD-explorer/explore/1.1/dev/)

(<http://arxiv.org/abs/1606.05250>)

### Getting Started

---

We've built a few resources to help you get started with the dataset.

Download a copy of the dataset (distributed under the CC BY-SA 4.0 (<http://creativecommons.org/licenses/by-sa/4.0/legalcode>) license):

[Training Set v2.0 \(40 MB\) \(/SQuAD-explorer/dataset/train-v2.0.json\)](/SQuAD-explorer/dataset/train-v2.0.json)

[Dev Set v2.0 \(4 MB\) \(/SQuAD-explorer/dataset/dev-v2.0.json\)](/SQuAD-explorer/dataset/dev-v2.0.json)

To evaluate your models, we have also made available the evaluation script we will use for official evaluation, along with a sample prediction file that the script will take as input. To run the evaluation, use `python evaluate-v2.0.py <path_to_dev-v2.0> <path_to_predictions>`.

[Evaluation Script v2.0](https://worksheets.codalab.org/rest/bundles/0x6b567e1cf2e041ec80d7098f031c5c9e/content)  
(<https://worksheets.codalab.org/rest/bundles/0x6b567e1cf2e041ec80d7098f031c5c9e/content>)

Sample Prediction File (on Dev v2.0)

(<https://worksheets.codalab.org/bundles/0x8731effab84f41b7b874a070e40f61e2/>)

Once you have built a model that works to your expectations on the dev set, you submit it to get official scores on the dev and a hidden test set. To preserve the integrity of test results, we do not release the test set to the public. Instead, we require you to submit your model so that we can run it on the test set for you. Here's a tutorial walking you through official evaluation of your model:

Submission Tutorial

(<https://worksheets.codalab.org/worksheets/0x8212d84ca41c4150b555a075b19ccc05/>)

Because SQuAD is an ongoing effort, we expect the dataset to evolve. To keep up to date with major changes to the dataset, please subscribe:

Subscribe

## Have Questions?

Ask us questions at our google group (<https://groups.google.com/forum/#!forum/squad-stanford-qa>) or at [pranavsr@stanford.edu](mailto:pranavsr@stanford.edu) (<mailto:pranavsr@stanford.edu>) and [robinjia@stanford.edu](mailto:robinjia@stanford.edu) (<mailto:robinjia@stanford.edu>).

## Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph. How will your system compare to humans on this task?

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> (Rajpurkar & Jia et al. '18) ( <a href="http://arxiv.org/abs/1606.05250">http://arxiv.org/abs/1606.05250</a> )	86.831	89.452
1 <span>Nov 16, 2018</span>	AoA + DA + BERT (ensemble) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	<b>82.374</b>	<b>85.310</b>
2 <span>Nov 16, 2018</span>	AoA + DA + BERT (single model) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	81.178	84.251
3 <span>Nov 16, 2018</span>	Candi-Net+BERT (single model) <i>42Maru NLP Team</i>	80.106	82.862
3 <span>Nov 08, 2018</span>	BERT (single model) <i>Google AI Language</i>	80.005	83.061
4 <span>Nov 09, 2018</span>	L6Net + BERT (single model) <i>Layer 6 AI</i>	79.181	82.259
5	SLQA+BERT (single model)	77.003	80.209

Nov 06, 2018	Alibaba DAMO NLP <a href="http://www.aclweb.org/anthology/P18-1158">http://www.aclweb.org/anthology/P18-1158</a> ( <a href="http://www.aclweb.org/anthology/P18-1158">http://www.aclweb.org/anthology/P18-1158</a> )		
6 Nov 08, 2018	BERT_base_aug (ensemble) <i>GammaLab</i>	76.721	79.611
7 Nov 05, 2018	MIR-MRC(F-Net) (single model) <i>Kangwon National University, Natural Language Processing Lab. &amp; ForceWin, KP Lab.</i>	74.791	77.988
8 Sep 13, 2018	nlnet (single model) <i>Microsoft Research Asia</i>	74.272	77.052
9 Oct 12, 2018	YARCS (ensemble) <i>IBM Research AI</i>	72.670	75.507
10 Oct 13, 2018	RNANetSimple (ensemble) <i>Anonymous</i>	72.580	75.075
11 Sep 17, 2018	Unet (ensemble) <i>Fudan University &amp; Liulishuo Lab</i> <a href="https://arxiv.org/abs/1810.06638">https://arxiv.org/abs/1810.06638</a> ( <a href="https://arxiv.org/abs/1810.06638">https://arxiv.org/abs/1810.06638</a> )	71.417	74.869
11 Nov 12, 2018	BASE(single model) <i>Hithink RoyalFlush</i>	71.496	74.815
12 Aug 28, 2018	SLQA+ (single model) <i>Alibaba DAMO NLP</i> <a href="http://www.aclweb.org/anthology/P18-1158">http://www.aclweb.org/anthology/P18-1158</a> ( <a href="http://www.aclweb.org/anthology/P18-1158">http://www.aclweb.org/anthology/P18-1158</a> )	71.462	74.434
12 Aug 15, 2018	Reinforced Mnemonic Reader + Answer Verifier (single model) <i>NUDT</i> <a href="https://arxiv.org/abs/1808.05759">https://arxiv.org/abs/1808.05759</a> ( <a href="https://arxiv.org/abs/1808.05759">https://arxiv.org/abs/1808.05759</a> )	71.767	74.295
12 Nov 14, 2018	BERT+Answer Verifier (single model) <i>Pingan Tech Olatop Lab</i>	71.666	75.457
13 Nov 15, 2018	PAML (ensemble) <i>GammaLab</i>	71.248	74.016
13 Sep 14, 2018	SAN (ensemble model) <i>Microsoft Business Applications AI Research</i> <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> ( <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> )	71.316	73.704
14 Oct 13, 2018	RNANetSimple (single model) <i>Anonymous</i>	70.718	73.403
15 Sep 14, 2018	Unet (single model) <i>Fudan University &amp; Liulishuo Lab</i>	69.262	72.642
15 Aug 21, 2018	FusionNet++ (ensemble) <i>Microsoft Business Applications Group AI Research</i> <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> ( <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> )	70.300	72.484

15 Sep 26, 2018	Multi-Level Attention Fusion(MLAF) (single model) <i>Chonbuk National University, Cognitive Computing Lab.</i>	69.476	72.857
16 Aug 21, 2018	SAN (single model) <i>Microsoft Business Applications AI Research</i> <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> ( <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> )	68.653	71.439
16 Aug 25, 2018	ARRR (single model) <i>anonymous</i>	68.653	71.124
16 Sep 13, 2018	BiDAF++ with pair2vec (single model) <i>UW and FAIR</i>	68.021	71.583
17 Jul 13, 2018	VS <sup>3</sup> -NET (single model) <i>Kangwon National University in South Korea</i>	67.897	70.884
17 Jun 24, 2018	KACTEIL-MRC(GFN-Net) (single model) <i>Kangwon National University, Natural Language Processing Lab.</i>	68.213	70.878
18 Jun 25, 2018	KakaoNet2 (single model) <i>Kakao NLP Team</i>	65.719	69.381
19 Jul 11, 2018	abcNet (single model) <i>Fudan University &amp; Liulishuo AI Lab</i>	65.256	69.206
19 Sep 13, 2018	BiDAF++ (single model) <i>UW and FAIR</i>	65.651	68.866
20 Jun 27, 2018	BSAE AddText (single model) <i>reciTAL.ai</i>	63.338	67.422
21 Aug 14, 2018	eeAttNet (single model) <i>BBD NLP Team</i> <a href="https://www.bbdservice.com">https://www.bbdservice.com</a> ( <a href="https://www.bbdservice.com">https://www.bbdservice.com</a> )	63.327	66.633
21 May 30, 2018	BiDAF + Self Attention + ELMo (single model) <i>Allen Institute for Artificial Intelligence [modified by Stanford]</i>	63.372	66.251
22 May 30, 2018	BiDAF + Self Attention (single model) <i>Allen Institute for Artificial Intelligence [modified by Stanford]</i>	59.332	62.305
23 May 30, 2018	BiDAF-No-Answer (single model) <i>University of Washington [modified by Stanford]</i>	59.174	62.093

## SQuAD1.1 Leaderboard

Since the release of SQuAD1.0, the community has made rapid progress, with the best models now rivaling human performance on the task. Here are the ExactMatch (EM) and F1 scores evaluated on the test set of v1.1.

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> (Rajpurkar et al. '16) ( <a href="http://arxiv.org/abs/1606.05250">http://arxiv.org/abs/1606.05250</a> )	82.304	91.221

1 Oct 05, 2018	BERT (ensemble) <i>Google AI Language</i> <a href="https://arxiv.org/abs/1810.04805">https://arxiv.org/abs/1810.04805</a> ( <a href="https://arxiv.org/abs/1810.04805">https://arxiv.org/abs/1810.04805</a> )	87.433	93.160
2 Oct 05, 2018	BERT (single model) <i>Google AI Language</i> <a href="https://arxiv.org/abs/1810.04805">https://arxiv.org/abs/1810.04805</a> ( <a href="https://arxiv.org/abs/1810.04805">https://arxiv.org/abs/1810.04805</a> )	85.083	91.835
2 Sep 09, 2018	nlnet (ensemble) <i>Microsoft Research Asia</i>	85.356	91.202
2 Sep 26, 2018	nlnet (ensemble) <i>Microsoft Research Asia</i>	85.954	91.677
3 Jul 11, 2018	QANet (ensemble) <i>Google Brain &amp; CMU</i>	84.454	90.490
4 Jul 08, 2018	r-net (ensemble) <i>Microsoft Research Asia</i>	84.003	90.147
5 Mar 19, 2018	QANet (ensemble) <i>Google Brain &amp; CMU</i>	83.877	89.737
5 Sep 09, 2018	nlnet (single model) <i>Microsoft Research Asia</i>	83.468	90.133
5 Jun 20, 2018	MARS (ensemble) <i>YUANFUDAO research NLP</i>	83.982	89.796
6 Sep 01, 2018	MARS (single model) <i>YUANFUDAO research NLP</i>	83.185	89.547
7 Jan 03, 2018	r-net+ (ensemble) <i>Microsoft Research Asia</i>	82.650	88.493
7 May 09, 2018	MARS (single model) <i>YUANFUDAO research NLP</i>	82.587	88.880
7 Feb 19, 2018	Reinforced Mnemonic Reader + A2D (ensemble model) <i>Microsoft Research Asia &amp; NUDT</i>	82.849	88.764
7 Jan 22, 2018	Hybrid AoA Reader (ensemble) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	82.482	89.281
7 Jun 20, 2018	QANet (single) <i>Google Brain &amp; CMU</i>	82.471	89.306
7 Mar 06, 2018	QANet (ensemble) <i>Google Brain &amp; CMU</i>	82.744	89.045
7 Jun 21, 2018	MARS (single model) <i>YUANFUDAO research NLP</i>	83.122	89.224
8 Jan 05, 2018	SLQA+ (ensemble) <i>Alibaba iDST NLP</i>	82.440	88.607
9 Feb 02, 2018	Reinforced Mnemonic Reader (ensemble model) <i>NUDT and Fudan University</i>	82.283	88.533

9 Feb 27, 2018	QANet (single model) <i>Google Brain &amp; CMU</i>	82.209	88.608
10 Dec 22, 2017	AttentionReader+ (ensemble) <i>Tencent DPDAC NLP</i>	81.790	88.163
11 May 09, 2018	Reinforced Mnemonic Reader + A2D (single model) <i>Microsoft Research Asia &amp; NUDT</i>	81.538	88.130
11 Dec 17, 2017	r-net (ensemble) <i>Microsoft Research Asia</i> <a href="http://aka.ms/rnet">http://aka.ms/rnet</a> ( <a href="http://aka.ms/rnet">http://aka.ms/rnet</a> )	82.136	88.126
12 May 09, 2018	Reinforced Mnemonic Reader + A2D + DA (single model) <i>Microsoft Research Asia &amp; NUDT</i>	81.401	88.122
12 Apr 23, 2018	r-net (single model) <i>Microsoft Research Asia</i>	81.391	88.170
12 Apr 03, 2018	KACTEIL-MRC(GF-Net+) (ensemble) <i>Kangwon National University, Natural Language Processing Lab.</i>	81.496	87.557
12 Feb 27, 2018	QANet (single model) <i>Google Brain &amp; CMU</i>	80.929	87.773
12 Nov 17, 2017	BiDAF + Self Attention + ELMo (ensemble) <i>Allen Institute for Artificial Intelligence</i>	81.003	87.432
12 Feb 19, 2018	Reinforced Mnemonic Reader + A2D (single model) <i>Microsoft Research Asia &amp; NUDT</i>	80.919	87.492
13 Feb 12, 2018	Reinforced Mnemonic Reader + A2D (single model) <i>Microsoft Research Asia &amp; NUDT</i>	80.489	87.454
13 Apr 12, 2018	AVIQA+ (ensemble) <i>aviqa team</i>	80.615	87.311
14 Mar 20, 2018	DNET (ensemble) <i>QA geeks</i>	80.164	86.721
14 Jan 22, 2018	Hybrid AoA Reader (single model) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	80.027	87.288
14 Jan 12, 2018	EAZI+ (ensemble) <i>Yiwise NLP Group</i>	80.426	86.912
14 Jan 13, 2018	SLQA+ <i>single model</i>	80.436	87.021
14 Jan 04, 2018	{EAZI} (ensemble) <i>Yiwise NLP Group</i>	80.436	86.912
15 Feb 12, 2018	BiDAF + Self Attention + ELMo + A2D (single model) <i>Microsoft Research Asia &amp; NUDT</i>	79.996	86.711
16 Jan 29, 2018	Reinforced Mnemonic Reader (single model) <i>NUDT and Fudan University</i>	79.545	86.654

<https://arxiv.org/abs/1705.02798>  
(<https://arxiv.org/abs/1705.02798>)

16 <a href="#">Apr 10, 2018</a>	Unnamed submission by null	80.027	86.612
16 <a href="#">Feb 23, 2018</a>	MAMCN+ (single model) <i>Samsung Research</i>	79.692	86.727
16 <a href="#">Jan 03, 2018</a>	r-net+ (single model) <i>Microsoft Research Asia</i>	79.901	86.536
16 <a href="#">Dec 28, 2017</a>	SLQA+ (single model) <i>Alibaba iDST NLP</i>	79.199	86.590
16 <a href="#">Dec 05, 2017</a>	SAN (ensemble model) <i>Microsoft Business AI Solutions Team</i> <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> ( <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> )	79.608	86.496
17 <a href="#">Oct 17, 2017</a>	Interactive AoA Reader+ (ensemble) <i>Joint Laboratory of HIT and iFLYTEK</i>	79.083	86.450
18 <a href="#">Oct 24, 2017</a>	FusionNet (ensemble) <i>Microsoft Business AI Solutions Team</i> <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> ( <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> )	78.978	86.016
18 <a href="#">Jun 01, 2018</a>	MDReader <i>single model</i>	79.031	86.006
18 <a href="#">Feb 01, 2018</a>	Unnamed submission by null	78.999	86.151
19 <a href="#">Oct 24, 2018</a>	WDNet (single model) <i>Beijing Normal University</i>	78.926	85.810
19 <a href="#">Oct 22, 2017</a>	DCN+ (ensemble) <i>Salesforce Research</i> <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> ( <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> )	78.852	85.996
20 <a href="#">Mar 29, 2018</a>	KACTEIL-MRC(GF-Net+) (single model) <i>Kangwon National University, Natural Language Processing Lab.</i>	78.664	85.780
20 <a href="#">Nov 03, 2017</a>	BiDAF + Self Attention + ELMo (single model) <i>Allen Institute for Artificial Intelligence</i>	78.580	85.833
21 <a href="#">May 09, 2018</a>	KakaoNet (single model) <i>Kakao NLP Team</i>	78.401	85.724
22 <a href="#">Nov 30, 2017</a>	SLQA(ensemble) <i>Alibaba iDST NLP</i>	78.328	85.682
22 <a href="#">Jan 02, 2018</a>	Conductor-net (ensemble) <i>CMU</i> <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> ( <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> )	78.433	85.517

22 Jun 01, 2018	MDReader0 <i>single model</i>	78.171	85.543
22 Sep 18, 2018	BiDAF++ with pair2vec (single model) <i>UW and FAIR</i>	78.223	85.535
22 Jan 03, 2018	MEMEN (single model) <i>Zhejiang University</i> <a href="https://arxiv.org/abs/1707.09098">https://arxiv.org/abs/1707.09098</a> ( <a href="https://arxiv.org/abs/1707.09098">https://arxiv.org/abs/1707.09098</a> )	78.234	85.344
22 Mar 19, 2018	aviqa (ensemble) <i>aviqa team</i>	78.496	85.469
23 Jan 29, 2018	test <i>single</i>	78.087	85.348
24 Jul 25, 2017	Interactive AoA Reader (ensemble) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	77.845	85.297
25 Jan 10, 2018	Unnamed submission by null	77.436	85.130
26 Dec 06, 2017	AttentionReader+ (single) <i>Tencent DPDAC NLP</i>	77.342	84.925
26 Sep 18, 2018	BiDAF++ (single model) <i>UW and FAIR</i>	77.573	84.858
26 Dec 13, 2017	RaSoR + TR + LM (single model) <i>Tel-Aviv University</i> <a href="https://arxiv.org/abs/1712.03609">https://arxiv.org/abs/1712.03609</a> ( <a href="https://arxiv.org/abs/1712.03609">https://arxiv.org/abs/1712.03609</a> )	77.583	84.163
26 Apr 10, 2018	Unnamed submission by null	77.489	84.735
26 Mar 20, 2018	DNET (single model) <i>QA geeks</i>	77.646	84.905
27 Nov 06, 2017	Conductor-net (ensemble) <i>CMU</i> <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> ( <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> )	76.996	84.630
27 Sep 26, 2018	{gqa} (single model) <i>FAIR</i>	77.090	83.931
27 Dec 21, 2017	Jenga (ensemble) <i>Facebook AI Research</i>	77.237	84.466
27 Jan 23, 2018	MARS (single model) <i>YUANFUDAO research NLP</i>	76.859	84.739
28 Nov 01, 2017	SAN (single model) <i>Microsoft Business AI Solutions Team</i> <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> ( <a href="https://arxiv.org/abs/1712.03556">https://arxiv.org/abs/1712.03556</a> )	76.828	84.396
29	r-net (single model)	76.461	84.265



Oct 13, 2017	Microsoft Research Asia <a href="http://aka.ms/rnet">http://aka.ms/rnet</a> ( <a href="http://aka.ms/rnet">http://aka.ms/rnet</a> )		
29 Dec 19, 2017	FRC (single model) <i>in review</i>	76.240	84.599
29 May 14, 2018	VS^3-NET (single model) <i>Kangwon National University in South Korea</i>	76.775	84.491
30 Oct 22, 2017	Conductor-net (ensemble) <i>CMU</i>	76.146	83.991
31 Sep 08, 2017	FusionNet (single model) <i>Microsoft Business AI Solutions team</i> <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> ( <a href="https://arxiv.org/abs/1711.07341">https://arxiv.org/abs/1711.07341</a> )	75.968	83.900
31 Oct 18, 2018	KAR (single model) <i>York University</i> <a href="https://arxiv.org/abs/1809.03449">https://arxiv.org/abs/1809.03449</a> ( <a href="https://arxiv.org/abs/1809.03449">https://arxiv.org/abs/1809.03449</a> )	76.125	83.538
32 Jul 14, 2017	smarnet (ensemble) <i>Eigen Technology &amp; Zhejiang University</i>	75.989	83.475
32 Oct 22, 2017	Interactive AoA Reader+ (single model) <i>Joint Laboratory of HIT and iFLYTEK</i>	75.821	83.843
32 Mar 15, 2018	AVIQA-v2 (single model) <i>aviqa team</i>	75.926	83.305
33 Oct 05, 2018	Unnamed submission by null	74.950	83.294
33 Aug 18, 2017	RaSoR + TR (single model) <i>Tel-Aviv University</i> <a href="https://arxiv.org/abs/1712.03609">https://arxiv.org/abs/1712.03609</a> ( <a href="https://arxiv.org/abs/1712.03609">https://arxiv.org/abs/1712.03609</a> )	75.789	83.261
34 Oct 23, 2017	DCN+ (single model) <i>Salesforce Research</i> <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> ( <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> )	75.087	83.081
35 Feb 13, 2018	SSR-BiDAF <i>ensemble model</i>	74.541	82.477
35 Nov 01, 2017	Mixed model (ensemble) <i>Sean</i>	75.265	82.769
36 Jan 02, 2018	Conductor-net (single model) <i>CMU</i> <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> ( <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> )	74.405	82.742
36 Nov 17, 2017	two-attention-self-attention (ensemble) <i>guotong1988</i>	75.223	82.716
36 May 21, 2017	MEMEN (ensemble) <i>Eigen Technology &amp; Zhejiang University</i>	75.370	82.658

<https://arxiv.org/abs/1707.09098>  
(<https://arxiv.org/abs/1707.09098>)

37 Mar 09, 2017	ReasoNet (ensemble) <i>MSR Redmond</i> <a href="https://arxiv.org/abs/1609.05284">https://arxiv.org/abs/1609.05284</a> ( <a href="https://arxiv.org/abs/1609.05284">https://arxiv.org/abs/1609.05284</a> )	75.034	82.552
38 Aug 14, 2018	eeAttNet (single model) <i>BBD NLP Team</i> <a href="https://www.bbdservice.com">https://www.bbdservice.com</a> ( <a href="https://www.bbdservice.com">https://www.bbdservice.com</a> )	74.604	82.501
38 Jul 10, 2017	DCN+ (single model) <i>Salesforce Research</i> <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> ( <a href="https://arxiv.org/abs/1711.00106">https://arxiv.org/abs/1711.00106</a> )	74.866	82.806
38 Feb 06, 2018	Jenga (single model) <i>Facebook AI Research</i>	74.373	82.845
38 Oct 27, 2017	Unnamed submission by null	74.489	82.312
38 Oct 31, 2017	SLQA (single model) <i>Alibaba iDST NLP</i>	74.489	82.815
39 Jul 14, 2017	Mnemonic Reader (ensemble) <i>NUDT and Fudan University</i> <a href="https://arxiv.org/abs/1705.02798">https://arxiv.org/abs/1705.02798</a> ( <a href="https://arxiv.org/abs/1705.02798">https://arxiv.org/abs/1705.02798</a> )	74.268	82.371
40 Dec 23, 2017	S <sup>3</sup> -Net (ensemble) <i>Kangwon National University in South Korea</i>	74.121	82.342
41 Jul 25, 2017	Interactive AoA Reader (single model) <i>Joint Laboratory of HIT and iFLYTEK Research</i>	73.639	81.931
41 Jul 29, 2017	SEDT (ensemble model) <i>CMU</i> <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> ( <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> )	74.090	81.761
42 Dec 14, 2017	Jenga (single model) <i>Facebook AI Research</i>	73.303	81.754
42 Nov 06, 2017	Conductor-net (single) <i>CMU</i> <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> ( <a href="https://arxiv.org/abs/1710.10504">https://arxiv.org/abs/1710.10504</a> )	73.240	81.933
42 Apr 22, 2017	SEDT+BiDAF (ensemble) <i>CMU</i> <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> ( <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> )	73.723	81.530
42 Feb 22, 2017	BiDAF (ensemble) <i>Allen Institute for AI &amp; University of Washington</i> <a href="https://arxiv.org/abs/1611.01603">https://arxiv.org/abs/1611.01603</a> ( <a href="https://arxiv.org/abs/1611.01603">https://arxiv.org/abs/1611.01603</a> )	73.744	81.525

42 Jan 24, 2017	Multi-Perspective Matching (ensemble) <i>IBM Research</i> <a href="https://arxiv.org/abs/1612.04211">https://arxiv.org/abs/1612.04211</a> ( <a href="https://arxiv.org/abs/1612.04211">https://arxiv.org/abs/1612.04211</a> )	73.765	81.257
42 Jul 06, 2017	SSAE (ensemble) <i>Tsinghua University</i>	74.080	81.665
43 May 01, 2017	jNet (ensemble) <i>USTC &amp; National Research Council Canada &amp; York University</i> <a href="https://arxiv.org/abs/1703.04617">https://arxiv.org/abs/1703.04617</a> ( <a href="https://arxiv.org/abs/1703.04617">https://arxiv.org/abs/1703.04617</a> )	73.010	81.517
44 Oct 22, 2017	Conductor-net (single) <i>CMU</i>	72.590	81.415
44 Apr 17, 2018	Unnamed submission by null	72.831	80.622
44 Nov 16, 2017	two-attention-self-attention (single model) <i>guotong1988</i>	72.600	81.011
44 Apr 12, 2017	T-gating (ensemble) <i>Peking University</i>	72.758	81.001
44 Sep 20, 2017	BiDAF + Self Attention (single model) <i>Allen Institute for Artificial Intelligence</i> <a href="https://arxiv.org/abs/1710.10723">https://arxiv.org/abs/1710.10723</a> ( <a href="https://arxiv.org/abs/1710.10723">https://arxiv.org/abs/1710.10723</a> )	72.139	81.048
45 Dec 15, 2017	S <sup>3</sup> -Net (single model) <i>Kangwon National University in South Korea</i>	71.908	81.023
45 Apr 17, 2018	Unnamed submission by null	72.831	80.622
46 Mar 03, 2018	AVIQA (single model) <i>aviqa team</i>	72.485	80.550
47 Nov 06, 2017	attention+self-attention (single model) <i>guotong1988</i>	71.698	80.462
48 Nov 01, 2016	Dynamic Coattention Networks (ensemble) <i>Salesforce Research</i> <a href="https://arxiv.org/abs/1611.01604">https://arxiv.org/abs/1611.01604</a> ( <a href="https://arxiv.org/abs/1611.01604">https://arxiv.org/abs/1611.01604</a> )	71.625	80.383
49 Jul 14, 2017	smarnet (single model) <i>Eigen Technology &amp; Zhejiang University</i> <a href="https://arxiv.org/abs/1710.02772">https://arxiv.org/abs/1710.02772</a> ( <a href="https://arxiv.org/abs/1710.02772">https://arxiv.org/abs/1710.02772</a> )	71.415	80.160
49 Apr 13, 2017	QFASE <i>NUS</i>	71.898	79.989
50 Oct 27, 2017	M-NET (single) <i>UFL</i>	71.016	79.835
50	MAMCN (single model)	70.985	79.939

Apr 22, 2018		<i>Samsung Research</i>		
50	Jul 14, 2017	Mnemonic Reader (single model) <i>NUDT and Fudan University</i> <a href="https://arxiv.org/abs/1705.02798">https://arxiv.org/abs/1705.02798</a> ( <a href="https://arxiv.org/abs/1705.02798">https://arxiv.org/abs/1705.02798</a> )	70.995	80.146
50	May 23, 2018	AttReader (single) <i>College of Computer &amp; Information Science, SouthWest University, Chongqing, China</i>	71.373	79.725
50	Mar 24, 2017	jNet (single model) <i>USTC &amp; National Research Council Canada &amp; York University</i> <a href="https://arxiv.org/abs/1703.04617">https://arxiv.org/abs/1703.04617</a> ( <a href="https://arxiv.org/abs/1703.04617">https://arxiv.org/abs/1703.04617</a> )	70.607	79.821
50	Apr 02, 2017	Ruminating Reader (single model) <i>New York University</i> <a href="https://arxiv.org/abs/1704.07415">https://arxiv.org/abs/1704.07415</a> ( <a href="https://arxiv.org/abs/1704.07415">https://arxiv.org/abs/1704.07415</a> )	70.639	79.456
50	Mar 14, 2017	Document Reader (single model) <i>Facebook AI Research</i> <a href="https://arxiv.org/abs/1704.00051">https://arxiv.org/abs/1704.00051</a> ( <a href="https://arxiv.org/abs/1704.00051">https://arxiv.org/abs/1704.00051</a> )	70.733	79.353
50	Dec 28, 2016	FastQAExt <i>German Research Center for Artificial Intelligence</i> <a href="https://arxiv.org/abs/1703.04816">https://arxiv.org/abs/1703.04816</a> ( <a href="https://arxiv.org/abs/1703.04816">https://arxiv.org/abs/1703.04816</a> )	70.849	78.857
50	May 13, 2017	RaSoR (single model) <i>Google NY, Tel-Aviv University</i> <a href="https://arxiv.org/abs/1611.01436">https://arxiv.org/abs/1611.01436</a> ( <a href="https://arxiv.org/abs/1611.01436">https://arxiv.org/abs/1611.01436</a> )	70.849	78.741
50	Mar 08, 2017	ReasoNet (single model) <i>MSR Redmond</i> <a href="https://arxiv.org/abs/1609.05284">https://arxiv.org/abs/1609.05284</a> ( <a href="https://arxiv.org/abs/1609.05284">https://arxiv.org/abs/1609.05284</a> )	70.555	79.364
51	Apr 14, 2017	Multi-Perspective Matching (single model) <i>IBM Research</i> <a href="https://arxiv.org/abs/1612.04211">https://arxiv.org/abs/1612.04211</a> ( <a href="https://arxiv.org/abs/1612.04211">https://arxiv.org/abs/1612.04211</a> )	70.387	78.784
52	Aug 30, 2017	SimpleBaseline (single model) <i>Technical University of Vienna</i>	69.600	78.236
52	Feb 05, 2018	SSR-BiDAF <i>single model</i>	69.443	78.358
53	Apr 12, 2017	SEDT+BiDAF (single model) <i>CMU</i> <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> ( <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> )	68.478	77.971
54	Jun 25, 2017	PQMN (single model) <i>KAIST &amp; AIBrain &amp; Crosscert</i>	68.331	77.783

55 Apr 12, 2017	T-gating (single model) <i>Peking University</i>	68.132	77.569
56 Nov 28, 2016	BiDAF (single model) <i>Allen Institute for AI &amp; University of Washington</i> <a href="https://arxiv.org/abs/1611.01603">https://arxiv.org/abs/1611.01603</a> ( <a href="https://arxiv.org/abs/1611.01603">https://arxiv.org/abs/1611.01603</a> )	67.974	77.323
56 Feb 22, 2018	Unnamed submission by null	68.478	77.220
57 Feb 22, 2018	Unnamed submission by null	68.425	77.077
57 Dec 28, 2016	FastQA <i>German Research Center for Artificial Intelligence</i> <a href="https://arxiv.org/abs/1703.04816">https://arxiv.org/abs/1703.04816</a> ( <a href="https://arxiv.org/abs/1703.04816">https://arxiv.org/abs/1703.04816</a> )	68.436	77.070
57 Jul 29, 2017	SEDT (single model) <i>CMU</i> <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> ( <a href="https://arxiv.org/abs/1703.00572">https://arxiv.org/abs/1703.00572</a> )	68.163	77.527
58 Oct 26, 2016	Match-LSTM with Ans-Ptr (Boundary) (ensemble) <i>Singapore Management University</i> <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> ( <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> )	67.901	77.022
58 Jan 22, 2018	FABIR (Single Model) <i>in review</i>	67.744	77.605
59 Sep 19, 2017	AllenNLP BiDAF (single model) <i>Allen Institute for AI</i> <a href="http://allennlp.org/">http://allennlp.org/</a> ( <a href="http://allennlp.org/">http://allennlp.org/</a> )	67.618	77.151
60 Feb 05, 2017	Iterative Co-attention Network <i>Fudan University</i>	67.502	76.786
61 Jan 03, 2018	newtest <i>single model</i>	66.527	75.787
61 Nov 01, 2016	Dynamic Coattention Networks (single model) <i>Salesforce Research</i> <a href="https://arxiv.org/abs/1611.01604">https://arxiv.org/abs/1611.01604</a> ( <a href="https://arxiv.org/abs/1611.01604">https://arxiv.org/abs/1611.01604</a> )	66.233	75.896
62 Feb 24, 2018	Unnamed submission by null	65.992	75.469
63 Jan 10, 2018	Unnamed submission by null	64.796	74.272
64 Dec 09, 2017	Unnamed submission by ravioncodalab	64.439	73.921
64 Oct 26, 2016	Match-LSTM with Bi-Ans-Ptr (Boundary) <i>Singapore Management University</i> <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> ( <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> )	64.744	73.743

65 Feb 19, 2017	Attentive CNN context with LSTM NLPR, CASIA	63.306	73.463
66 Nov 02, 2016	Fine-Grained Gating Carnegie Mellon University <a href="https://arxiv.org/abs/1611.01724">https://arxiv.org/abs/1611.01724</a> ( <a href="https://arxiv.org/abs/1611.01724">https://arxiv.org/abs/1611.01724</a> )	62.446	73.327
66 Sep 21, 2017	OTF dict+spelling (single) University of Montreal <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> ( <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> )	64.083	73.056
67 Sep 21, 2017	OTF spelling (single) University of Montreal <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> ( <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> )	62.897	72.016
68 Sep 21, 2017	OTF spelling+lemma (single) University of Montreal <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> ( <a href="https://arxiv.org/abs/1706.00286">https://arxiv.org/abs/1706.00286</a> )	62.604	71.968
69 Sep 28, 2016	Dynamic Chunk Reader IBM <a href="https://arxiv.org/abs/1610.09996">https://arxiv.org/abs/1610.09996</a> ( <a href="https://arxiv.org/abs/1610.09996">https://arxiv.org/abs/1610.09996</a> )	62.499	70.956
70 Aug 27, 2016	Match-LSTM with Ans-Ptr (Boundary) Singapore Management University <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> ( <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> )	60.474	70.695
71 Sep 11, 2018	Unnamed submission by Will_Wu	59.058	69.436
72 Jan 10, 2018	Unnamed submission by null	58.764	69.276
73 Aug 27, 2016	Match-LSTM with Ans-Ptr (Sentence) Singapore Management University <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> ( <a href="https://arxiv.org/abs/1608.07905">https://arxiv.org/abs/1608.07905</a> )	54.505	67.748
74 Oct 26, 2018	Unnamed submission by minjoon	52.533	62.757