Math23K for Math Word Problem Solving

1. Introduction

Math23K is a dataset created for math word problem solving, contains 23, 162 **Chinese** problems crawled from internet.

1. Data Description

|  |  |  |
| --- | --- | --- |
| Field | Description | Example |
| id | Unique id of a problem. | 8 |
| original\_text | Original problem text before word segmentation | 买一套住房，售价28万元，每年付2万元，多少年后才能付清？ |
| Segmented\_text | problem text after word segmentation | 买 一套 住房 ， 售价 28 万元 ， 每年 付 2 万元 ， 多少 年 后 才能 付清 |
| equation | The equation used to solve the problem. | x=28/2 |
| ans | Problem answer | 14 |

Please cite our [EMNLP 2017 paper](http://www.aclweb.org/anthology/P/P16/P16-1084.pdf) [1] when you use this dataset in your research.

Reference:

[1] Wang, Yan, Xiaojiang Liu, and Shuming Shi. "Deep Neural Solver for Math Word Problems." *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017.

[2] Lei Wang, Yan Wang, Deng Cai, Dongxiang Zhang, Xiaojiang Liu. "Deep Neural Solver for Math Word Problems." *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. 2017.