

China's schools are quietly using AI to mark students' essays ... but do the robots make the grade?

China has made artificial intelligence-enabled education a national strategy. SCMP Graphics

One in every four schools in China is quietly testing a powerful machine that uses artificial intelligence to mark pupils' work, according to scientists involved in the government programme.

The technology is designed to understand the general logic and meaning of the text and make a reasonable, human-like judgment about the essay's overall quality.

It then grades the work, adding recommended improvements in areas such as writing style, structure and theme.

The technology, which is being used in around 60,000 schools, is supposed to "think" more deeply and do more than a standard spellchecker.

For instance, if a paragraph starts trailing off topic, the computer would mark it down.



The AI grading technology is supposed to help, rather than replace teachers. Photo: AFP

Scientists insist the technology is designed to assist, rather than replace, human teachers.

It could help to reduce the amount of time teachers spend on grading essays and help them avoid inconsistencies caused by human errors such as lapses in attention or unconscious bias.

It could also help more students, especially those in remote areas with limited access to resources, improve their writing skills more quickly.

The machine is similar to the e-rater, an automated system used by the Education Testing Service in the US to grade prospective postgraduate students' essays.

But unlike the e-rater, it can read both Chinese and English.

[China looks to school kids to win the global AI race](#)

Artificial intelligence is developing rapidly in China with strong support from the government and the technology is used in many areas of everyday life.

But the extensive tests of the essay grading machine – built by some of the leading language processing teams involved in the government and military's internet surveillance programme – were carried out with unusual security measures in place.

In most of the schools taking part in the programme, parents were not informed, access to the system terminals was limited to authorised staff, test results were strictly classified, and in some classes even the pupils were unaware that their work had been read and scored by a machine.



Schools are using the software for internal tests, rather than important national exams. Photo: AFP

Wang Jing, director of academic affairs office in the High School Affiliated to Renmin University, one of the country's most prestigious schools, said: "We are treating [the test] with extreme caution.

"What happens on campus stays on campus. The test results will not be revealed to the public," he added, in line with the school's agreement with the project organisers.

Most schools interviewed by the *South China Morning Post* – including the Baita Middle School in Nanchang, Sichuan province; the Fifth High School in Fuyang, Anhui and the 58th High School in Qingdao, Shandong – gave a similar assessment.

The schools said the AI grading machine was far from perfect, with teachers citing many examples where a brilliant piece of writing was given low marks.

The software is presently being used to mark only internal tests and none of the schools had plans to use the technology to grade essays in exams that would affect pupils' official academic record.

"It's still in its infancy," Wang said.

[How 'China's MIT' drives the country's technology ambitions](#)

But the developers say the machine is already 10 years old and they are increasingly confident about its potential.

A scientist involved in the project at the school of computer science and engineering at Beihang University in Beijing compared it to the AlphaGo, an AI Go player developed by Google which has defeated human world champions over the past couple of years.

The essay grading machine, embedded in a cluster of fast computers in Beijing, is improving its ability to understand human language by using deep learning algorithms to plough through essays written by Chinese students and "compare notes" with human teachers' grading and comments.

It is also able to collect and build its own "knowledge base" with little or no human intervention.

"It has evolved continuously and become so complex, we no longer know for sure what it was thinking and how it made a judgment," said the researcher, who requested not to be named due to the sensitivity of the project.



Chinese Go player Ke Jie in a contest against AlphaGo, which has beaten the world's top players. Photo: Chinatopix via AP

According to a government document seen by the *South China Morning Post*, the tests involved 60,000 schools with more than 120 million people involved.

The AI and human grader gave the same score 92 per cent of the time, but the document did not specify the content and scale of the tests.

The researcher confirmed the figures but declined to reveal more details.

"In the future it may be used to relieve the teacher's burden but it will never replace teachers. The machine has no soul," he added.

[Microsoft's AI bot writes poetry, hosts a TV show and can converse like a human in Mandarin](#)

The essay grading machine project was led by professor Zhou Jianshe, director of the research centre for language intelligence in China in Capital Normal University.

Zhou and other senior members of the project have received government and military awards for their contributions to natural language processing and mining information from big data.

Zhou could not be reached for comment.

The machine can be accessed from various online portals but are only open to registered users.

One English portal, pigai.org, requires a user to register either as a teacher or student and provide information, such as school name and class number, to verify their identity.

[This Chinese weapons lab is putting missile technology into airport body scanners](#)

Users gave a similarly mixed response to the machines. While some said they were useful and more accurate than similar essay grading systems overseas, others described them as rubbish.

Some users have argued the software cannot distinguish between academic essays and other forms of writing.

One user on Zhihu, the largest question-and-answer website in China, posted a screenshot showing how the machine had assessed an April 2015 *Washington Post* comment piece "Why is Obama sticking it to stay-at-home moms?" as if it were answering an essay question.

The piece got a score of 71.5 out of 100 and the machine said that while the vocabulary used was "rich and appropriate" it was "slightly short for academic language".

It concluded: "The flow can be improved on smoothness; and please improve the focus of the article; the paragraphs and sentences should be related to the topic."



句酷批改网
www.pigai.org

帮您轻松改作文

作文内容

In a speech in Rhode Island in October, President Obama offered his views about child care in a society where many parents have to make difficult choices: "Sometimes, someone, usually mom, leaves the workplace to stay home with the kids, which then leaves her earning a lower wage for the rest of her life as a result. And that's not a choice we want Americans to make."

I guess I made a choice we don't want Americans to make.

Thirty years ago, when my first child was born, I left my job as a lawyer. I stayed home with my daughter, and then my son, until they were in middle school. Then I started a second career, one in which I would no longer earn as much as my husband. That, I know, made me a contributor to the wage gap. But I have no regrets. My full-time stint at home, engaging my children and watching them thrive, was the most challenging and rewarding period of my life.

☐ 清空 ☐ 检查谓语动词 ☐ 检查语态一致 ☐ 知识点定制

句酷批改

分数评语

71.5

词汇:

句子:

篇章结构:

内容相关:

评语: 文中词汇表达较丰富恰当, 但学术词汇足; 文中应增加从句和复合句的使用; 行文稍应适当增加连接词的使用; 请提高文章内敛性句子要围绕主题展开。

A Washington Post article was graded by a machine – that could not tell that it was a newspaper opinion piece rather than an academic essay. Photo: Zhihu.com

Zhu Xiaoyan, head of the state key laboratory of intelligent technology and systems at Tsinghua University, said human language AI technology has achieved significant progress in recent years.

She said some machines have written articles that went viral on social media, attracting more than 10 million views, but did not provide further details.

But Zhu said she had not heard of the essay-grading program adding that she would not use a machine to grade her students' papers, adding: "It's a human job."

Yu Yafeng, professor at the institute of educational theories at Beijing Normal University, said computers could help grade candidates in subjects such as mathematics and physics because the answers were objective.

[Chinese surveillance programme mines data from workers' brains](#)

But essays can contain cultural, emotional or personal elements that a machine would not be able to gauge.

"There is no law forbidding AI from grading student essays, but this practice should raise ethical questions," she said.

An eight-year-old primary school pupil in Chaoyang district, Beijing said he did not mind an AI machine checking his essays, pointing out that his teachers already used readily available technology to check the answers to basic maths questions.

"Our teachers are already using mobile phones to grade our maths homework," he said.

"Take a photo and the score is out. What's the difference to an essay?"

This article appeared in the South China Morning Post print edition as: schools quietly using A.I. to mark student essays