ThinkCERCA A and B

In 2017, ThinkCERCA created a patented technology designed to enable a research-based approach to improving literacy development on a large scale.

Using a six-step process that helps students break the task of close reading, critical thinking, and academic writing across disciplines into smaller cognitive chunks, the technology walks students through a rigorous process resulting in a formal piece of writing.

One of the key types of writing that ThinkCERCA focuses on is argumentative writing. This genre works across grade levels and subject areas, and it is the most powerful type of writing in terms of college and career readiness. Not only does Thinkcerca help students improve their argumentative writing as proof of their learning, it also improves learning itself. It has been proven to enable students to achieve 2 years of reading growth per year and gain 20% growth in writing in as little as 8 weeks.

One of the key features of this process is the interactive argument-builder, or CERCA framework interaction, where students input argument components which often end up being part of their final piece. They can then organize those pieces in the argument-builder step or later as they are developing final drafts.

The CERCA framework has the following components:
Claim
Evidence
Reasoning
Counterargument
Audience

Over the course of 2021, ThinkCerca has been working with a variety of experts, learning scientists, data scientists, and engineers to enable machine learning that draws on the decades of student essays and teacher feedback captured in this technology. **One of these partnerships included collaborating with HDAG throughout the spring (with 2 case teams working on automated student essay and teacher feedback) and summer.**

Through a blend of inputs coming from the content woven into the process by ThinkCERCA's editorial team (prompts, texts, sentence stems, etc.,) student inputs in the six-step process, including the interactive, color-coded text boxes of the CERCA framework, and other sources of NLP, ThinkCERCA's team has developed a relatively reliable way to detect the components of the arguments students write. **The current model for identifying argument components is a fine-tuned model built with BERT, a powerful NLP pre-training tool developed by Google.** We have even done some testing on other texts outside of our own student essays.

There are several potential use cases for the algorithm that we've delineated based on the need for reliable, enabling feedback for readers, whether they are students, teachers, or others interested in reading and evaluating arguments.

Feedback tools that can supplement real-time feedback tools focused on punctuation, grammar, and syntax. Paired with the already excellent sentence-level tools available in the market, the feedback tools we're working on would provide substantive feedback during or after the drafting or formal writing process. The feedback tools we are developing can be used with, for, and by students or teachers.

Case Team Descriptions

Both projects - ThinkCerca has traditionally been very engaged with the work of HDAG case teams during partnerships. This entails weekly to biweekly meetings to update the client on progress with the CEO and other ThinkCerca stakeholders, with a collaborative back-and-forth partnership from week to week. Additionally, ThinkCerca has previously hired HDAG case team members for summer internships in data science!

ThinkCerca A - This team will focus on building a model for evaluating the quality of different aspects of writing (e.g. claim, evidence, reason, etc.), building upon previous work by HDAG and ThinkCerca to tag different parts of each argument. Conceptually, this will entail mapping out key aspects of "good writing" such as cohesion, concision, sentence structure variety, etc. Technically, this will involve heavy usage of natural language processing in order to analyze qualitative factors at the sentence/word level quantitatively, as well as machine learning / data analytics to work with existing ThinkCerca assignment data to test the predictive power of these ideas.

Emphasized Skills for Team A: Natural Language Processing, Data Analytics, Machine Learning/Modeling, Visualization

2) ThinkCerca B - This team will focus on building a model for socially responsible NLP, applied to evaluating student writing. This project will involve a significant conceptual and research aspect into different social/environmental factors which affect writing syntax, and how to mitigate biases in NLP methods as applied to ThinkCerca's model for evaluating student writing. Technically, this would likely heavily involve natural language processing in order to glean out different patterns of communication across a variety of school districts and other factors. Additionally, the team could work to create a metric for evaluating the context of a particular district/school/class/student in order to more individually tailor feedback.

Emphasized Skills for Team B: Natural Language Processing, Data Analytics, Machine Learning/Modeling, Research