Natural Language Processing for Automated Evaluation of Student Writing

Jill Burstein and Claudia Leacock, Educational Testing Service

There is an increased use of educational applications based on natural language processing for both large-scale assessment and classroom instruction. This has occurred for two primary reasons. First, there has been a significant increase in the availability of computers in schools. Second, there has been notable development in computer-based educational applications that incorporate advanced methods in natural language processing that can be used to evaluate students’ work.

Applications that automate the evaluation of essays use text-based NLP techniques that employ feature extraction and statistical machine learning. Such applications are commercially available, and are being used internationally by thousands of students. We will discuss these writing evaluation capabilities in the context of a deployed application, CriterionSM Online Essay Evaluation Service. Criterion has two complementary applications. The first, e-rater, is an automated essay scoring system. The second, Critique Writing Analysis Tools, is a suite of programs that detect errors in grammar, usage, and mechanics. The two applications identify discourse elements in the essay and recognize elements of undesirable style.

Dr. Jill Burstein and Dr. Claudia Leacock work at Educational Testing Service in Princeton.