

Outomatiese assessering van eerstejaar HTML- en CSS- werkopdragte

J (Jocelyne) Smith, E Kotzé, L Nel

Departement Rekenaarwetenskap en Informatika, Universiteit van die Vrystaat, Suid-Afrika

Korresponderende outeur: Jocelyne Smith **E-pos:** me@jocelyne.co.za

Automatic assessment of first-year HTML and CSS assignments: This study explores how AI, specifically large language models (LLMs), can enhance automated grading for introductory computer science courses. The research develops an AI-driven grading system, compares local and API-based LLMs, and finds significant reductions in grading time (up to 97%). It highlights the potential of AI to improve grading efficiency and scalability.

Dosente ervaar uitdagings om koderingswerkopdragte vinnig en akkuraat na te sien weens 'n skerp toename in studentinskrywings. Hierdie studie ondersoek die gebruik van kunsmatige intelligensie, spesifiek groot taalmodelle, vir die outomatiese nasien van HTML- en CSS-werkopdragte onder eerstejaarstudente. Die navorsing bestaan uit twee fases, naamlik eerstens die ontwikkeling van 'n outomatiese nasienstelsel, en tweedens die evaluering van groot taalmodelle.

Die eerste fase behels die ontwikkeling van 'n outomatiese nasienstelsel wat groot taalmodelle gebruik om memorandum met studente-werkopdragte te vergelyk. Hierdie modelle assessee die opdragte en verskaf individuele terugvoer. Die stelsel maak gebruik van optiese karakterherkenning om vrae uit die memorandum te onttrek en om as konteks vir gradering te gebruik. Verder gebruik die studie gevorderde KI-tegnieke om die akkuraatheid en terugvoer te verbeter. Hierdie tegnieke sluit in:

- Refleksie – 'n iteratiewe tegniek wat 'n groot taalmodel toelaat om terugvoer te verbeter op grond van vorige evalueringe.
- Hulpmiddelgebruik – die integrasie van verskeie KI-nutsprogramme.
- Multi-agent samewerking – die gebruik van 'n gespesialiseerde agent vir elke taak (byvoorbeeld gradering van werkopdragte, onttrekking van vrae uit die memorandum en lees van die inhoud van werkopdragte), met besprekings tussen agente.

Die tweede fase vergelyk plaaslike oopbronmodelle (byvoorbeeld Llama-2 en Gemma-2) met toepassingsprogrammeerkoppelvlakmodelle (TPK-modelle, byvoorbeeld ChatGPT en Claude). Die studie se doelwit is om te bepaal watter model die beste balans van doeltreffendheid, koste, akkuraatheid, nasientyd en bruikbaarheid bied. Die navorsing volg die grondslagdata wetenskapmetodologie, wat prosesse soos besigheidbegrip, modellering, evaluering en implementering insluit. Hierdie benadering verseker die stelselmatige ontwikkeling van die nasienstelsel.

Daar is bevind dat die outomatiese nasienstelsel die spoed van gradering aansienlik verbeter het. Die evaluering dui op 'n afname in nasientyd van ongeveer 97%. Gevorderde KI-tegnieke soos refleksie, hulpmiddelgebruik, en multi-agent samewerking het die akkuraatheid en relevansie van gradering beduidend verbeter. TPK-modelle lewer beter resultate aangesien dit kragtige grootskaalse infrastruktuur benut waartoe gebruikers nie gewoonlik toegang het nie.

Plaaslike oopbronmodelle soos Llama-2 en Gemma-2 het goeie resultate gelewer maar die proses was stadig omdat eindgebruikers nie noodwendig toegang tot die mees moderne hardeware het nie. Die ChatGPT-modelle verskaf die mees gedetailleerde en akkurate graderings. ChatGPT-4-mini is die model van keuse, aangesien dit 'n balans bied tussen akkuraatheid, nasientyd en koste-effektiwiteit.

Die gevolgtrekking is dat hierdie navorsing die kritieke behoefte aan skaalbare en akkurate nasienoplossings te midde van die stygende getal inskrywings van universiteitstudente aanspreek. Die breër implikasies van hierdie navorsing dui daarop dat sulke stelsels aangepas kan word om opvoeders in verskeie kontekste by te staan en moontlik KI in onderrig aan die Universiteit van die Vrystaat bekend te stel. Toekomstige werk kan fokus op die verfyning van die stelsel gebaseer op gebruikersterugvoer en die uitbreiding van die toepassing in bykomende areas, soos verskillende koderingstale en -raamwerke.

Nota: 'n Seleksie van referaatopsommings: Studentesimposium in die Natuurwetenskappe, 30-31 Oktober 2024, Universiteit van die Vrystaat. Reëlingskomitee: Prof Rudi Pretorius (Departement Geografie, Universiteit van Suid-Afrika); Dr Hertzog Bisset (Suid-Afrikaanse Kernenergie-korporasie); Dr Ernie Langner (Departement Chemie, Universiteit van die Vrystaat); Dr Wynand Nel (Departement Rekenaarwetenskap en Informatika, Universiteit van die Vrystaat) en Prof Liesl van As (Departement Dierkunde en Entomologie, Universiteit van die Vrystaat).