ANALYTICAL APPROACHES FOR ANALYSIS OF STUDENT TEXTS WITHOUT AN EVALUATION FRAMEWORK

Iskra Petkova
Medical University – Pleven, Bulgaria, petkovai@abv.bg
Tsvetomir Petkov
Bulgarian Institute of Metrology – Sofia, Bulgaria, petkov_@abv.bg
Danelina Vacheva
Thracian University – Stara Zagora, Bulgaria, danelina@dbv.bg

Abstract: This article examines different analytical behaviors of language models and their influence on the way student texts are interpreted and used in an educational environment. The aim is to present the integrative possibilities of artificial intelligence for structural-logical and factual analysis through the use of a Standard GPT-5.2 model (Standard Model) and an author-developed logical scaffold, LOGIC_CRUTCH_v2025.291225_BALKAN_PARADOX_ALPHA (Algorithm 291225). Materials and Methods: An author-developed logical crutch (scaffold), 291225_BALKAN_PARADOX_ALPHA, created within ChatGPT and based on the GPT-5.2 model, is used. An analysis of written academic tasks is conducted through the integrated use of the Standard Model and Algorithm 291225 in order to formulate a final assessment and provide guidance for working with the student. Results: Academic texts from two groups of written learning tasks (speech disorders; motor disabilities) for independent preparation of students are used as empirical material for research through a consistent integrated application and comparison of the Standard Model and Algorithm 291225. Lecture materials from the respective discipline are used as the primary source in the students' work. The Standard Model performs: a detailed analysis of the completed tasks for each student by comparing them with the accuracy of scientific data; identifies critical errors (factual errors and omissions in the student's answers) where present; proposes corrections and provides an assessment of completeness and factual correctness. The Standard Model closes the logical chain by striving to produce an "acceptable," "correct," and "complete" answer. Algorithm 291225 identifies critical errors as paradoxes, which are not forcibly closed but are mapped against the lecture material (accepted as the source of the expected "truth") without being corrected; builds a logical bridge that determines the presence/absence of metadata as well as their complete/incomplete information; provides concrete feedback to the student and structures guidance for further work. The feedback to the student is directed toward discovering authorial thinking as a value, even when it leads to factual errors. The sequential use of the Standard Model and Algorithm 291225 enables the application of an integrated strategy to the completion of assigned academic tasks, resulting in both an objective assessment and a deep understanding of the student's cognitive process. Conclusion: The practical utility of such an approach is manifested mostly in working with the learning process, rather than with the final result. Algorithm 291225 can function as a tool to support reflection – both on the part of the student and the teacher – without automating or replacing pedagogical judgment. In this sense, the distinction between the Standard Model and Algorithm 291225 provides not ready-made solutions, but a framework for a more conscious and responsible use of language technologies in education. Recommendation: The development of a university project to study the possibilities of applying a strategy for the integrated use of the Standard Model and Algorithm 291225 in support of students' cognitive abilities in higher education.
Keywords: artificial intelligence, standard model, algorithm, integrated strategy.

1. INTRODUCTION

The entry of language models into educational practice poses questions for teachers that go beyond their technical use and concern the very logic of academic assessment (Luckin et al., 2016). In the context of working with student texts, these tools are increasingly used as means of analysis, feedback, and writing support, but their application often reproduces traditional evaluative frameworks based on correction, classification, and compliance with predefined norms. The current opinion proceeds from the understanding that the analytical potential of language models is not exhausted by automated evaluation of results, but can be used to conceptualize the process of writing and thinking itself (Vasileva, 2024).

Instead of positioning the model as an arbiter of „correctness", the analysis considers the possibility that it may function as a tool for explicating structural choices, argumentative strategies, and internal tensions in student texts. In this sense, the distinction between assessment and sense-making is not presented as an opposition, but as a difference in analytical focus (Kirilov, 2024). Whereas the evaluative approach presupposes hierarchization and closure of interpretation, the sense-making approach preserves the openness of the text and admits contradictions,

ambiguities, and incompleteness as part of the learning process (Dimitrova, 2024). This opens space for pedagogical feedback that does not aim at immediate correction, but supports the development of analytical and reflective writing (Holgaard et al., 2022).

Within this context, is of interest how different analytical behaviors of language models influence the way student texts are interpreted and used in the educational environment (Huang et al., 2023). The emphasis is not on the quality of the generated analysis, but on the logic it follows and on the pedagogical consequences of its application. In this way, is offered proposes a framework for using language models that consciously distances itself from the evaluative function and places the process of sense-making at the center (Ivanov and Mateva, 2024).

All this made determines the need to conduct an empirical study that **aims** to present the integrative possibilities of artificial intelligence (AI) for structural-logical and factual analysis when using the Standard Model GPT-5.2 and the authorial logical scaffold LOGIC_CRUTCH_v2025.291225 _BALKAN_PARADOX_ALPHA.

2. MATERIALS AND METHODS

An authorial development of a logical crutch (scaffold) is used – 291225_BALKAN_PARADOX_ALPHA (Algorithm 291225), created in ChatGPT, based on model GPT-5.2 (Standard Model), with Knowledge Cutoff: August 2025, on the basis of Legacy Protocol LOGIC_CRUTCH 2025 (v. April Update). The analyses are carried out under the conditions of a free ChatGPT account, where the main model is also GPT-5.2, but with Knowledge Cutoff: April 2024 (to date) and Legacy Protocol LOGIC_CRUTCH_2024 (Stabilized). The aim is to use the crutch also in free accounts, as its functionality has been demonstrated (it was created) for future upgrades of ChatGPT.

An analysis of written academic tasks is conducted through the integrated use of the Standard Model and Algorithm 291225 in order to formulate a final evaluation and provide guidance for working with the student.

3. RESULTS

The study was conducted among two groups of students (8 students each) from the specialty „Social Activities in Healthcare“, Faculty of Public Health, Medical University – Pleven, Bulgaria. The students were assigned tasks for self-preparation in the course „Fundamentals of Special Pedagogy“. The educational tasks are structured in two thematic groups:

- Speech disorders – „comparative characteristics of voice disorders“ and „comparative characteristics of articulation disorders“;
- Motor impairments – „comparative characteristics of some of the most common physical impairments in children – (similarities and differences) between the conditions Cerebral Palsy (ICD-10, G80) and Poliomyelitis (sequelae of Poliomyelitis, ICD-10, B91)“ and „characteristics of the disease Rickets (ICD-10, E55.0) – features, possibilities for education“.

Academic texts from two groups of written self-preparation tasks are used as empirical material for the study through successive integrated application and comparison of the Standard Model and Algorithm 291225.

In this case, the *Standard Model (mode: Contextual Mapping)* performs:

- a detailed *analysis* of the completed tasks for each student, comparing them with the accuracy of scientific data. For example, in the work of Ts. V. the analysis shows: „The answer is extremely detailed and factually correct (p. 1). The causes (brain damage vs. virus) (p. 1), the mechanisms of transmission (none vs. airborne/fecal-oral) (p. 1), and prevention (no vaccine vs. mandatory vaccine) (p. 2) are correctly distinguished“. In the answer to the second question by D. B., for example, partial incompleteness is observed: „The definition is short and correct – a general disease related to metabolic disorders (p. 1). Rickets is a metabolic disease related to vitamin D deficiency. Information about features and possibilities for education is missing“;
- identifies *critical errors* (factual errors and omissions in the student's answers) where present – e.g., D. B.: „the claim that both (CP and Poliomyelitis) are congenital diseases“ (p. 1 of the paper); N. R.: the claim „Even if they have good intellectual potential, those suffering from this disease have difficulties in learning and usually cannot achieve great success“ (p. 2);
- proposes *correction* – e.g., for the paper of N. R.: „Rickets caused by vitamin D deficiency affects physical but not intellectual development of the child, except in very rare cases, etc.“;
- performs an *assessment* of exhaustiveness and factual correctness. For example, for Ts. V.: „The answer is complete, exhaustive, and covers all aspects of the task,“ and for D. B.: „The answer is minimalist and incomplete according to the requirements of the task“.

Typically here, the Standard Model closes the logical chain because it strives to produce an „acceptable“, „correct“, and „complete“ answer.

Use of Algorithm 291225 (mode: Paradox Handling):

- critical errors are defined as *paradoxes*. They are not forcibly closed, but mapped against the lecture material (accepted as the source of the expected „truth“) without being corrected. For example, in the paper of D. B. the paradoxes are structured as:

Paradox 1 (Congenital vs. Infectious): There is a direct contradiction between the student's claim („both are congenital“ (p. 1)) and the medical facts (Poliomyelitis is infectious, not congenital). Logic Bridge suggests that the student may be confusing „childhood“ (occur in childhood) with „congenital“ (present at birth). This is an illogicality that is not smoothed out but registered as incorrect conceptual understanding.
Paradox 2 (Rickets): The answer is minimalist and incomplete according to the task requirements. Contextual Mapping with the lecture material would mark the lack of detail on education and features, probably focusing on the medical definition.

- Algorithm 291225 builds a *logical bridge* that determines the presence/absence of metadata and their completeness/incompleteness. For example, in the paper of Ts. V., the Logic Bridge shows that the material is full of *metadata* (complete information on ICD-10, etiology, clinical picture, prevention) (pp. 1–2), which does not require supplementation. The answer demonstrates structural truth;
- specific *feedback* is provided to the student, combining objective assessment with analysis of the logical process and structuring of guidelines for work. Feedback is directed at discovering authorial thinking as a value, even when it leads to factual errors. For example, for D. B. and N. R., *the guidelines* focus on encouraging the logical process while correcting the facts. The main goal is to use their natural seeking toward logical construction (*authorial thinking*) and channel it toward fact-checking and avoidance of logical fallacies.

The use of the Standard Model and Algorithm 291225 in sequence makes it possible to obtain both objective evaluation and in-depth understanding of the student's cognitive process. The process goes through three stages (Table 1).

Table 1: Strategy for integrated use of the Standard Model and Algorithm 291225

Stage 1: Application of the Standard Model (Product Assessment)	*Stage 2: Application of Algorithm 291225 (Process Analysis)*	*Stage 3: Synthesis and Final Assessment*
Begins with this algorithm for rapid and objective evaluation.	Builds upon the results from Stage 1 with this algorithm for qualitative and descriptive analysis.	Combines the two perspectives for balanced feedback.
Aim: Rapid marking of correctness of the answer against lecture material and medical facts.	*Aim:* Understanding why the student made the error (mapping the paradox) and assessing the level of authorial thinking.	*Integration:* Uses the objective assessment (Stage 1) and qualitative analysis (Stage 2) to determine the final grade and provide targeted feedback. *Example*: „The answer is factually incorrect (Standard Model), but the analysis shows that the error stems from a logical (though incorrect) interpretation of a social context (Algorithm 291225), which indicates authorial thinking.“
Action: Identify all „critical errors“, „omissions“ and „incompleteness“. This provides quantitative and qualitative assessment of the student's product.	*Action:* Take the errors marked by the standard algorithm („paradoxes“) and use Logic Bridge and Contextual Mapping to hypothesize the logical path or confusion that led to them. Assess the presence of "structural truth" or missing "metadata" (e.g., lack of pedagogical aspect).	
Result: Clear evaluation whether the answer is „correct“ or „incorrect“ and where exactly the problems are.	*Result:* In-depth insight into the student's thinking process.	

Source: Author research from consistent use and comparison of Standard Model and Algorithm 291225

For the present publication, the application of an integrated strategy to the performance of the assigned tasks by students is of interest. For example, the integrated analysis of the task of N. R. shows the following results:
Stage 1: Application of the Standard Model (Product Assessment) – The answer is evaluated against objective medical and pedagogical facts:

- *Part 1* (CP and Poliomyelitis): Evaluated as excellent. All facts are accurate and complete, without errors.
- *Part 2* (Rickets): Evaluated as incorrect in a key statement.

- *Result of Stage 1*: The work is of high quality in Part 1 but contains a serious factual error in Part 2, making it partially incorrect.

Stage 2: Application of Algorithm 291225 (Process Analysis) – Builds upon this by analyzing the logic behind the error:

- *Part 1* (CP and Poliomyelitis): No paradoxes. Full structural truth and rich metadata.
- *Part 2* (Rickets): Structural paradox present.
- *Logic Bridge* (Explanation of the paradox): The student made a personal interpretation by equating physical deformities and potential social isolation with intellectual impairment. She assumed that physical illness automatically leads to cognitive difficulties. This is an example of authorial thinking (i.e., not copying), which, however, leads to an incorrect conclusion because metadata from the lecture material that explicitly deny this connection are missing (as in Z. R.).
- *Result of Stage 2*: The error is not random or careless, but the result of a personal, though incorrect, logical construction by the student.

Stage 3: Synthesis and Final Assessment – The synthesis combines the objective evaluation and the process analysis:

- The work of N. R. demonstrates excellent factual preparation on the topic of CP and Poliomyelitis (Stages 1 and 2). In the second part, however, a critical factual error is made that does not correspond to the academic consensus.
- *The key conclusion* is that this error is an indication of *the presence of authorial thinking and an attempt at independent logical construction* (Stage 2), rather than copying or using AI. The student made a logical error in interpreting the causal relationship between physical illness and intellect.
- *Final evaluation*: The work is mixed, showing potential for independent thinking but also a need for stricter adherence to the factual material provided in the lecture content.

This integrated evaluation provides a much more complete picture than that presented independently by each of the Standard Model and Algorithm 291225.

It is noteworthy that in the analysis of the first group of tasks on speech disorders related to „comparative characteristics of voice disorders" and „comparative characteristics of articulation disorders", paradoxes are observed arising from the overlay of two classification models: the Medical Model and the Linguistic-Behavioral Model. The possibility that paradoxes arise from the overlap of a medical and a pedagogical definition is extremely high and fully explains the inconsistencies found. This is a perfect example of Contextual Mapping in action. Here is how Algorithm 291225 (BALKAN_PARADOX_ALPHA) explains this overlay:

„Dual Interpretation: Medical vs. Pedagogical – Different paradigms use different classification criteria:

- *Medical definition* (etiology and pathology) – focuses on the cause (etiology) and the pathological mechanism;
- Pedagogical definition (correction and symptomatology) – focuses on the symptom, the corrective process, and the child's age.

In this case, the paradox arises because both truths are valid – the medical classification is ontologically more precise, while the pedagogical one is methodologically more practical".

The choice, and in some cases simplification, of the model (the content of the concept) is conditionally by the specifics of training in the professional field and specialty. All this is determined by the teacher within the Curriculum and in the development of interdisciplinary scientific content, where classifications from two or more scientific fields are considered (e.g., medicine and pedagogy/speech therapy).

4. DISCUSSION

It is essential to seek and differentiate the differences and specificities in the use of the Standard Model and Algorithm 291225, as well as the possibilities for their integrated use in determining final evaluation and balanced feedback.

The main difference lies in the philosophy of information processing and the aim of the analysis. The distinction between the Standard Model and Algorithm 291225 is not considered as an assessment of quality, but as a difference in the epistemological function that both perform in the educational context. The analysis aims to make these differences visible in order to support their more conscious use in pedagogical practice, without reducing them to tools for automated assessment or control. The underlying logic of both the Standard Model and Algorithm 291225 should be understood as a result of their way of processing information, not as a predefined pedagogical strategy approaching problem-based learning, an issue addressed by Holgaard, J.E. et al. (2022). The Standard Model is oriented toward stability and consensus, striving to reduce inconsistencies and present the text as a finished result. In

contrast, Algorithm 291225 explicates thinking processes, including incompleteness and internal tensions, which leads to a different type of analytical behavior and different pedagogical applicability.

This difference is clearly manifested in the attitude toward the text. When the text is viewed as a finished object, the focus naturally turns to its correctness and conformity to expectations. In the integrated approach, the text is interpreted as an indicator of cognitive dynamics, allowing the analysis to concentrate on argumentation, logical transitions, and the choice of frameworks without necessarily seeking final „correctness". On the issue of perception and evaluation of text from an interview with a real person versus an interview with AI, Despina Vasileva (2024) notes consolidated answers when working with an „artificial" interview, with statistically significant differences in favor of AI in two out of ten questions and one in favor of the real interview. In this context, contradictions and paradoxes are not treated as defects but as carriers of analytical information. The Standard Model tends to minimize such tensions through generalization or reformulation, whereas Algorithm 291225 preserves them as active elements of analysis. This opens the possibility for pedagogical use of contradictions—as a starting point for discussion and development rather than as grounds for automatic correction.

Nikola Kirilov et al. (2024) investigate the use of ChatGPT for student self-preparation and find that 32.43% of respondents already know and use this chatbot. A significant part (75%) of those who use it fully trust ChatGPT without checking the authenticity of generated texts, which can negatively affect learning and skills formation. Despite the difficulties, students demonstrate a desire to become familiar with and increasingly use digital technologies, including AI, in their learning. The introduction of regulatory frameworks could optimize the role of ChatGPT in education, fully confirming our conclusions about the need for analytical thinking when perceiving information from AI. Similar conclusions are reached by Nelly Dimitrova (2024), who points out advantages and disadvantages of AI in student learning.

The attitude toward assessment also differs substantially. The lack of explicit evaluative markers in Algorithm 291225 does not mean a refusal of pedagogical position, but a shift of focus from classification to sense-making. Instead of defining the text as „good" or „bad", the analysis makes visible the structural choices behind it, allowing a form of feedback oriented toward development without imposing hierarchy. In this direction is the White Paper on AI in Europe (https://eur-lex.europa.), seeking high performance and an atmosphere of trust by formulating a people-centered, ethical, inclusive, and secure approach to AI development and use, respecting the values with which the EU wishes to associate. On this basis, the Concept for the Development of Artificial Intelligence in Bulgaria until 2030 has been developed (https://eur-lex.europa.eu/).

The distinction between different types of hallucinations should be understood as an analytical observation, not as a strict typology. While in the Standard Model additions often create a false sense of completeness and stability, in Algorithm 291225 analytical tension remains visible and therefore more easily recognizable. In an educational context, this supports a more critical attitude toward both the student's text and the tool itself.

Research on developing student thinking through experimental learning and inclusion of AI tools conducted by Fawad Ahmed et al. (2025) emphasizes the limitations of traditional, theory-heavy teaching approaches and confirms our conclusions that AI-enhanced learning experiences create more sustainable changes in thinking by overcoming the critical gap between abstract knowledge and practical application. The study expands current understanding by showing how generative AI tools serve as cognitive scaffolds that amplify the learning cycle, especially in the phases of reflection and active experimentation, while maintaining human oversight for ethical engagement (Tsvetkova, 2025).

The widely widespread claim that most AI systems do not assess collaboration, social skills, and creative abilities (Ethical Guidelines, 2022) is based on understanding assessment as a final, classifying operation. The introduction of analytical models such as Algorithm 291225 problematizes this assumption by shifting the focus from assessing abilities to analyzing their structural manifestations in the text. Instead of treating these qualities as missing objects of assessment, it makes them visible through analysis of their structural manifestations, thus revealing that the problem is not the absence of assessment, but the reductive understanding of what „assessment" means in an educational context.

5. CONCLUSION

The practical utility of such an approach is manifested mostly in working with the learning process, rather than with the final result. Algorithm 291225 can function as a tool to support reflection – both on the part of the student and the teacher – without automating or replacing pedagogical judgment. In this sense, the distinction between the Standard Model and Algorithm 291225 provides not ready-made solutions, but a framework for a more conscious and responsible use of language technologies in education.

RECOMMENDATION

The development of a university project to study the possibilities of applying a strategy for the integrated use of the Standard Model and Algorithm 291225 in support of students' cognitive abilities in higher education.

REFERENCES

Ahmed, F., Ying, T.L., & Shu, H.C. (2025). Developing Entrepreneurial Mindset Among Non-Business Majors Through Experiential Learning and AI Tools. *Proceedings of the 26th European Conference* on Knowledge Management, Vol. 26, (1):1-7. https://doi.org/10.34190/eckm.26.1.3762

Dimitrova, N.St. (2024). Predimstva i nedostatatsi na izkustvenia intelekt pri obuchenieto na studentite. *Godishnik na Shumenski universitet* „Episkop Konstantin Preslavski". Universitetsko izdatelstvo – Shumen, god. 28, (1):702-705. doi:10.1007/s11423-023-10188-5 [In Bulgarian]

Evropeyska komisiya. (2022). Byala kniga za izkustveniya intelekt – Evropa v tarsene na visoki postizheniya i atmosfera na doverie. Sluzhba za publikatsii na Evropeyskiya sayuz, Bryuksel.
https://op.europa.eu/bg/publication-detail/-/publication/ac957f13-53c6-11ea-aece-01aa75ed71a1 [accessed January 19, 2026]. [In Bulgarian]

Evropeyska komisiya. (2022). Etichni nasoki za prepodavatelite otnosno izpolzvaneto na izkustven intelekt (II) i na danni pri prepodavane i uchene (II). Oblast 4: Otsenyavane. Izpolzvane na tsifrovi tehnologii i strategii za podobryavane na otsenyavaneto. Sluzhba za publikatsii na Evropeyskiya sayuz, Lyuksemburg, p. 30.
https://op.europa.eu/bg/publication-detail/-/publication/d81a0d54-5348-11ed-92ed-01aa75ed71a1/language-en [accessed January 19, 2026]. [In Bulgarian]

Holgaard, J.E., Du, X. & Guerra, A. (2022). Problem-based learning in engineering education and its impact on developing an entrepreneurial mindset. *European Journal of Engineering Education*, 47(1):24-39.

Huang, R.H., Liu, D.J., Tlili, A., Yang, J.F. & Wang, H.H. (2023). Generative AI in higher education: Opportunities and challenges. *Educational Technology Research and Development*, 71(1):1-6.

Ivanov, I. & Mateva, T. (2024). Naglasi otnosno izpolzvaneto na izkustvenia intelekt v uchebnia protses. *Nauchni trudove na Kolezh* – Dobrich kam Shumenski universitet „Episkop Konstantin Preslavski". Universitetsko izdatelstvo – Shumen, god. 28, 2024, (1):160. [In Bulgarian]

Kirilov, N., Bischoff, E., Kovachev, M., Todorova, C., Vladeva, S. (2024). Prouchvane izpolzvaneto na CHATGPT za samopodgotovka na studentite. Study the Use of CHATGPT for Student Self-Training. *Sbornik dokladi ot nauchna konferentsia* „Znanie, nauka, inovatsii, tehnologii". Vol. 1, (2): 554-563. [In Bulgarian]

Luckin, R., Holmes, W., Griffiths, M. & Forcier, L.B. (2016). Intelligence unleashed: An argument for AI in education. London: Pearson Education.

Ministerstvo na transporta, informatsionnite tehnologii i saobshteniyata. (2020). Kontseptsiya za razvitieto na izkustveniya intelekt v Balgariya do 2030 g. Izkustven intelekt za inteligenten rastezh i prosperirashto demokratichno obshtestvo. Sofiya.
https://www.mtc.government.bg/sites/default/files/koncepciyazarazvitienaiivbulgariyado2030.pdf [accessed January 19, 2026]. [In Bulgarian]

Tsvetkova, D. (2025). Izpolzvaneto na digitalni instrumenti v universitetskoto obuchenie kato faktor za profesionalnata realizatsia na studentite. *Industrialni otnoshenia i obshtestveno razvitie*, (3). [In Bulgarian]

Vasileva, D. (2024). Perception and Evaluation of Text from an Interview with a Real Person and an Interview with an Artificial Intelligence. *Chuzhdoezikovo Obuchenie-Foreign Language Teaching journalis.* Vol. 51, (3):304-313. https://www.ceeol.com/search/article-detail?id=1272391 [accessed January 19, 2026].