

Εργασία 8^η

Διονύσης Κοτζαΐτης

Μελετήστε την εργασία

Sidney D'Mello, Nia Dowell, and Art Graesser. 2009. Cohesion Relationships in Tutorial Dialogue as Predictors of Affective States. In Proceedings of the 2009 conference on Artificial Intelligence in Education: Building Learning Systems that Care: From Knowledge Representation to Affective Modelling. IOS Press, NLD, 9–16

και απαντήστε τις παρακάτω ερωτήσεις. (Αποφεύγετε να αντιγράφετε σε μεγάλο βαθμό αυτούσια τμήματα της εργασίας ως απαντήσεις στις ερωτήσεις. Οι απαντήσεις να είναι ακριβείς, περιεκτικές. Αποφεύγετε αυτόματες μεταφράσεις όπως από Google translate).

Introduction

1. Ποια είναι τα πλεονεκτήματα από τη χρήση συστημάτων που επιτρέπουν τη εκτίμηση της συναισθηματικής κατάστασης (affective state) μέσω διαλόγων σε Tutorials;

Τα πλεονεκτήματα είναι:

- Μπορούμε να λάβουμε πληθώρα φθηνών feature μέσω των έξυπνων συστημάτων διδασκαλίας, τα οποία χρησιμοποιούν διαλόγους φυσικής γλώσσας.
- Οι «γραφές» (textual features) προήλθαν από διαλόγους διδασκαλίας είναι περιορισμένες όσον αφορά το context με τρόπο που γίνεται νύξη ως προς την κοινωνική πολιτική του δασκάλου και του μαθητή.
- Μπορούν να προβλέψουν πιο σύνθετα φαινόμενα όπως προσωπικότητα, εξαπάτηση ή/και θέματα υγείας.

2. Τα υπάρχοντα μοντέλα που βασίζονται στην επεξεργασία φυσικής γλώσσας πως προσπαθούν να αντιληφθούν το συναίσθημα του χρήστη. Αυτή η προσέγγιση μπορεί να εφαρμοστεί στον εκπαιδευτικό χώρο; Ναι ή όχι και γιατί;

Τα δεδομένα αυτά χρησιμοποιούνται μεν για να προβλέψουν τα συναισθήματα του χρήστη, παρόλα αυτά δεν ενδείκνυνται για την χρήση στον εκπαιδευτικό χώρο, καθώς:

- Ενώ χρησιμοποιούνται για την μελέτη συναισθημάτων κομματιών κειμένου, δεν λαμβάνουν υπόψιν ότι ο συναισθηματικός τόνος βρίσκεται αποκλειστικά μέσα στο γραπτό, κάτι που δεν είναι σίγουρο ότι ισχύει στα tutors, αλλά ισχύει το εντελώς αντίθετο.
- Οι μαθητές έχουν συναισθήματα κατά την διάρκεια των απαντήσεων, τα οποία όμως δεν εκφράζονται μέσα από τις απαντήσεις στα tutors, όπου οι απαντήσεις είναι domain specific και άμεσες. Έτσι, ίσως χρειαστεί μια πιο λεπτομερή μελέτη του κειμένου.

3. Στην μελέτη αυτή πως διαπιστώθηκε η συναισθηματική κατάσταση των εκπαιδευόμενων;

Η διαπίστωση έγινε με βάση τις σχέσεις cohesion που υπάρχουν σε φυσικούς διαλόγους του λογισμικού. Πρακτικά, η κατάσταση μετρήθηκε με την συνοχή του κειμένου, το οποίο είναι ένα μετρήσιμο χαρακτηριστικό και οδηγεί στην ψυχολογική κατάσταση του μαθητή.

4. Από τη ψυχολογία τι γνωρίζουμε σχετικά με τη συνοχή του γραπτού λόγου (textual cohesion) και για τη συναισθηματική κατάσταση;

Το cohesion από την ψυχολογία ξέρουμε ότι:

- Συνδέεται με το coherence.
- Συνδέεται με τη νοητική αναπαράσταση των ουσιαστικών ιδεών του αναγνώστη.
- Μπορούν να προβλεφθούν οι συναισθηματικές εμπειρίες των μαθητών.
- Διαταραχή συνεκτικότητας → σύγχυση και απογοήτευση
- Σταθερή συνεκτικότητα → συγκέντρωση.

Methods

5. Δώστε πληροφορίες για τους συμμετέχοντες, γνωστικό αντικείμενο, υλικό (ερωτήσεις, κτλ) αυτής της μελέτης.

Οι συμμετέχοντες ήταν:

- 28 προπτυχιακοί φοιτητές, με «έπαθλο» μια επιπλέον πιστωτική μονάδα.
- Πανεπιστήμιο στο Mid-South της Αμερικής.
- Κάθε μαθητής αλληλοεπίδρασε με το AutoTutor λογισμικό, για 32 λεπτά σε τυχαίο θέμα στην επιστήμη των υπολογιστών.

6. Ποιο λογισμικό χρησιμοποιήθηκε; περιγράψτε πως είναι οργανωμένοι οι διάλογοι στο λογισμικό αυτό.

Το λογισμικό που χρησιμοποιήθηκε είναι το AutoTutor. Οι διάλογοι είναι οργανωμένοι γύρω από δύσκολες ερωτήσεις και προβλήματα τα οποία χρειάζονται λογική και επεξήγηση στις απαντήσεις. Χρειάζονται 3-7 προτάσεις περίπου ανα απάντηση σε φυσική γλώσσα, αν και δίνονται συνήθως απαντήσεις 1-2 γραμμών. Γι αυτόν το λόγο το σύστημα δίνει ώθηση στους μαθητές σε έναν διάλογο που βοηθάει τον μαθητή στο να μάθει και να δώσει ολοκληρωμένες απαντήσεις.

7. Κατά τη διάρκεια της χρήσης του λογισμικού τι καταγράφηκε;

- Βίντεο του προσώπου των συμμετεχόντων.
- Μοτίβο στάσης σώματος.
- Βίντεο οθόνης.

8. Η καταγραφή των συναισθηματικών καταστάσεων των συμμετεχόντων πως έγινε και από ποιούς;

Για την καταγραφή των συναισθημάτων κάθε χρήστη έγινε από ένα σετ 4 «κριτών»:

- Ο ίδιος ο μαθητής, αμέσως μετά τις απαντήσεις του.
- Οι συμμαθητές του, καθένας εκ των οποίων «ήρθε» μια εβδομάδα μετά για να κρίνει τις αντιδράσεις των συμμαθητών του.
- Δυο ειδικούς κριτές, οι οποίοι ήταν προπτυχιακοί φοιτητές οι οποίοι εκπαιδεύτηκαν μέσω του Ekman's Facial Action Coding System (FACS).

9. Ποιες συναισθηματικές καταστάσεις καταγράφηκαν; Γιατί επιλέχθηκαν αυτές;

- Boredom.
- Flow.
- Engagement.
- Σύγχυση.
- Απογοήτευση.
- Χαρά.
- Έκπληξη.
- Neutral.

Η επιλογή αυτών των συναισθηματικών καταστάσεων έγινε, γιατί ήταν τα κυρίαρχα συναισθήματα που παρατηρήθηκαν στους συμμετέχοντες σε προηγούμενες μελέτες που έγιναν με βάση το λογισμικό AutoTutor και με βάση άλλα περιβάλλοντα εκπαίδευσης.

10. Πόσες καταγραφές έγιναν από κάθε κατηγορία κριτή;

- self judgments → 2.967 καταγραφές
- peer judgments → 3.012 καταγραφές
- 2 trained judges → 2.995 καταγραφές από τον έναν κριτή και 3.093 καταγραφές από τον άλλον κριτή

11. Πως κατηγοριοποιήθηκαν οι διάλογοι και γιατί;

Οι διάλογοι κατηγοριοποιήθηκαν σε 2 κατηγορίες:

- Το ένα σύνολο αποτελείται από τα λεγόμενα των μαθητών κατά τις απαντήσεις και τον διάλογο τους με το καθηγητή, που δείχνουν έμμεσα την συναισθηματική τους κατάσταση.
- Το δεύτερο σύνολο επικεντρώνεται στα λεγόμενα του tutor στο διάλογο με τον μαθητή. Αυτά υποδεικνύουν γεγονότα που προκαλούν αλλαγές στη συναισθηματική κατάσταση των εκπαιδευόμενων ή την διατηρούν.

Γενικά, οι απαντήσεις των μαθητών θεωρούμε ότι προέρχονται από τις συναισθηματικές τους καταστάσεις τα οποία με την σειρά τους επηρεάζουν τις απαντήσεις του tutor.

Measures of Cohesion

12. Ποιο λογισμικό χρησιμοποιήθηκε για τη συνοχή των διαλόγων; Πόσους διαφορετικούς τύπους μέτρων υποστηρίζει το λογισμικό αυτό;

Χρησιμοποιήθηκε το coh-metrix program για να μετρήσουμε το cohesion των διαλόγων. Αυτό, μετράει πάνω από 100 μετρικές διαφορετικών τύπων.

13. Σύντομη περιγραφή των παρακάτω δεικτών συνοχής διαλόγων

1. Co-reference cohesion (συν-αναφορά συνοχής)

Ο Coh-Metrix βρίσκει το ποσοστό των γειτονικών προτάσεων που βρίσκονται σε ένα window 2 ή 3 προτάσεων που παρουσιάζουν επικάλυψη ενός ουσιαστικού και ενός stem. Χρησιμοποιείται για όταν ένα μέρος του λόγου υπάρχει σε γειτονικές προτάσεις.

2. Pronoun referential (αναφορική συνοχή των αντωνυμιών)

Χρησιμοποιείται όταν μία αντωνυμία εμφανίζεται σε μια πρόταση και σε μια αμέσως προηγούμενη πρόταση. Στο coh-metrix χρησιμοποιείται ένας προσεγγιστικός αλγόριθμος.

3. Causal cohesion (αιτιακή συνοχή)

Όταν μέρη του κειμένου που αναφέρονται σε δράσεις και γεγονότα συνδέονται με αιτιολογικούς συνδέσμους. Ο Coh-Metrix παρέχει μέτρα σχετικά με τη συχνότητα εμφάνισης αυτών των κατηγοριών του ρήματος που εκφράζουν ενέργειες και των αιτιολογικών συνδέσμων, για παράδειγμα παρέχει τη συχνότητα εμφάνισης των ρημάτων και αιτιολογικών συνδέσμων ανά 1000 λέξεις. Επιπλέον, υπολογίζεται η συχνότητα εμφάνισης των αιτιολογικών συνδέσμων αυτών σε σημεία του κειμένου.

4. *Semantic cohesion (σημασιολογική συνοχή)*

Μέσω αυτού μετράμε την συνοχή μεταξύ δυο γειτονικών μερών του κειμένου. Το coh-metrix μετράει την διαφορά ανάμεσα σε μέρη του κειμένου, χρησιμοποιώντας διαφορετικού μεγέθους παράθυρα. Το μέτρο που χρησιμοποιείται συνήθως επι του θέματος είναι το LSA που είναι όμοιο με cosine similarity και χρησιμοποιείται για ομοιότητα κειμένων.

5. *Connectives (Συνδετικές λέξεις)*

Μέρη του κειμένου που συνδέουν το νόημα. Γενικά σύνδεσμοι διαφόρων ειδών, με τον Coh να βαθμολογεί τις λέξεις.

6. *Shallow Measures of Readability and Verbosity (Μικρά μέτρα αναγνωσιμότητας και πολυλογίας/ομιλίας)*

Αποτελούνται από:

- Λέξεις μιας πρότασης.
- Τμήματα διαλόγου.
- Συνδετικές λέξεις.

14. *Οι διάλογοι για κάθε πρόβλημα πως υποβλήθηκαν στο Coh-Matrix;*

Οι διάλογοι προστέθηκαν ως ανεξάρτητα δεδομένα στο coh, χρησιμοποιώντας συγκεκριμένα επιπλέον χαρακτηριστικά- παράγοντες πρόβλεψης (25 σε αριθμό) που προέκυψαν απ τον διάλογο. Για κάθε predictor υπολογίζουμε μια βαθμολογία από το σύνολο των διαλόγων.

15. *Πόσοι διακριτοί δείκτες συγκεντρώθηκαν; Για κάθε δείκτη συνοχής διαλόγου πόσα μέτρα (measures) χρησιμοποιήσαν; Υπάρχει τεκμηρίωση για την επιλογή αυτών των μέτρων;*

Χρησιμοποιήθηκαν 12 δείκτες με συγκεκριμένο αριθμό μέτρων:

- 5 διακριτοί
 - Co-reference cohesion → 4 μέτρα
 - Pronoun referential cohesion → 1 μέτρο
 - Causal cohesion → 3 μέτρα
 - Semantic cohesion → 5 μέτρα
 - Connectives → 2 μέτρα
- 7 άλλοι δείκτες.

Η επιλογή των δεικτών δικαιολογήθηκε μέσω αποτελεσμάτων άλλων μελετών και βασίζεται στο αν μπορεί να γίνει διάκριση ανάμεσα σε υψηλό και χαμηλό cohesion.

Results and Discussion

16. Δώστε το ποσοστό για κάθε συναισθηματική κατάσταση που καταγράφηκε από τους κριτές;

- Neutral → 31.2 %
- Confusion → 23.9%
- Flow → 16.1%
- Boredom → 16.7%
- Απογοήτευση → 7.2%
- Χαρά → 3.2%
- Έκπληξη → 1.7%

17. Για ποιες συναισθηματικές καταστάσεις τελικώς κατασκεύασαν μοντέλα πρόβλεψης και γιατί;

Οι συγγραφείς εν τέλει χρησιμοποίησαν τις 5 πρώτες καταστάσεις σε ποσοστό για να κατασκευάσουν το μοντέλο πρόβλεψης καθώς αυτές κυριάρχησαν.

18. Διατυπώστε τους προβληματισμούς των συγγραφέων αναφορικά με τη εσωτερική αξιοπιστία (interrater reliability) των μετρήσεων καθώς συμμετείχαν 4 διαφορετικοί κριτές.

Αρχικά πρέπει να αναφέρουμε ότι το πρόβλημα εκπαιδεύεται και λύνεται με βάση μοντέλων παλινδρόμησης. Σ' αυτά οι 4 κριτές μπορούν να δημιουργήσουν επιπλοκές, καθώς:

- Θέμα με την εσωτερική αξιοπιστία. Δηλαδή Χ, Υ, Ζ διαφορετικά διαγνωστικά του συναισθήματος Ε ενός ατόμου όταν το μετράμε από το ίδιο άτομο, από peers και από τους ειδικούς αντίστοιχα.
- Διαφορετικά χαρακτηριστικά (Χ, Υ, Ζ) που προβλέπουν το ίδιο συναίσθημα όταν μετρώνται από διαφορετικούς κριτές υποθέτουν ότι η σχέση μεταξύ των χαρακτηριστικών και του συναισθήματος Ε εξαρτώνται από την επίδραση των κριτών.

19. Τελικώς πόσοι δείκτες χρησιμοποιήθηκαν συνολικά και ανά συναισθηματική κατάσταση και γιατί;

Οι συγγραφείς εν τέλει χρησιμοποίησαν 9 δείκτες :

- Boredom → 2.
- Confusion → 2
- Flow → 2.
- Απογοήτευση → 3

Αφορά τα χαρακτηριστικά που έχουν σημαντική συσχέτιση με 2 τουλάχιστον ratings κριτών κατάστασης προκειμένου να έχει νόημα η περεταίρω διερεύνηση τους.

20. Τι μοντέλα κατασκευάστηκαν; Ποιες οι ανεξάρτητες μεταβλητές και ποια η εξαρτημένη;

Χρησιμοποιήθηκε πολλαπλή παλινδρόμηση για κάθε συναίσθημα.

- Εξαρτημένη μεταβλητή: Για κάθε ανάλυση χρησιμοποιούμε την αναλογία εμφάνισης του συναισθήματος παίρνοντας τον μ.ο. των 4 κριτών.
- Ανεξάρτητες μεταβλητές: 2 ή 3 από τους διαγνωστικούς προβλεπτικούς δείκτες για κάθε συναίσθημα (αναφέρονται στο 19 ο αριθμός).
- Γίνεται παλινδρόμηση είτε με backwards είτε με forwards addition μεταβλητών, ώστε να απομονώσουμε μόνο τις μεταβλητές που θέλουμε.

21. Στον Table 1 τι παρουσιάζουν; Εξηγήστε τα αποτελέσματα κάθε γραμμής. Τι συμπεράσματα προκύπτουν για κάθε μια από τις 4 συναισθηματικές καταστάσεις;

Στον πίνακα βλέπουμε τις συσχετίσεις ανάμεσα στα 4 κυρίαρχα συναισθήματα και α) τις παραμέτρους των παλινδρομήσεων που χρησιμοποιούνται ως μοντέλα, β) τους προβλεπτικούς παράγοντες για τα συναισθήματα είτε αυτοί είναι βασισμένη στο διάλογο από την πλευρά του μαθητή, είτε από την πλευρά του tutor.

Πιο συγκεκριμένα, στις γραμμές έχουμε:

- F-score: Μεταβλητή η οποία ελέγχει αν τουλάχιστον ένας από τους κανόνες πρόβλεψης έχει στατιστικά σημαντική συσχέτιση με την αντίστοιχη στον πίνακα μεταβλητή. Εδώ βλέπουμε ότι μεγαλύτερο f-score έχει η μεταβλητή Flow, ενώ ακολουθείτε από τις boredom και frustration.
- Df1,df2: Οι βαθμοί ελευθερίας για τον έλεγχο υποθέσεων για το f-score. 3/4 τιμές έχουν 1.24 βαθμούς ελευθερίας, με την boredom να έχει 1.25
- P: το p-value που χρησιμοποιείται στον έλεγχο υποθέσεων. Αν η p-τιμή είναι μικρότερη του 0.05 σημαίνει ότι τουλάχιστον μία από τις ανεξάρτητες μεταβλητές συσχετίζεται σημαντικά με ένα συναίσθημα (εξαρτημένη μεταβλητή). Όλες οι τιμές των συναισθημάτων είναι μικρότερες του 5% που σημαίνει ότι τουλάχιστον μια από τις ανεξάρτητες μεταβλητές είναι στατιστικά εξαρτημένη από το συναίσθημα.
- Adj. R Sq.: Το adjusted R² είναι μια τροποποιημένη μορφή του R² που εκτός από το να εκφράζει τη προσαρμογή του μοντέλου στα δεδομένα, λαμβάνει υπόψη το πλήθος των predictors του μοντέλου. Όσο περισσότερες μη σημαντικές μεταβλητές τόσο μικρότερο το adj 2 sq. Η μεγαλύτερη από τις τιμές είναι στο flow που δείχνει ότι ίσως έχουμε στο μοντέλο πρόβλεψης του κάποιες τιμές που δεν χρειάζονται τόσο.
- Standardized Coefficient weights: Κανονικοποιημένες μεταβλητές της παλινδρόμησης, με τέτοιο τρόπο ώστε standard deviation =1 . Οι συντελεστές αυτοί δείχνουν πόσες τυπικές αποκλίσεις θα μεταβληθεί η εξαρτημένη μεταβλητή, όταν γίνει αύξηση της ανεξάρτητης μεταβλητής κατά μία τυπική απόκλιση.

Στο β τμήμα του πίνακα έχουμε:

Model Coefficients	Συναίσθημα	Τιμή	Ερμηνεία
Incident of negations από τον Student	Boredom	.525 > 0	Η τιμή της μεταβλητής (θετική) δείχνει ότι όταν οι μαθητές βαριούνται κατά την διάρκεια της άσκησης θα προσθέσουν στις απαντήσεις τους πολλές φορές δεν ξέρω πχ.
Pronoun referential cohesion από tutor	Confusion	-.505 < 0	Εδώ ο αρνητικός συντελεστής της εξαρτημένης μεταβλητής των αντωνυμιών, δείχνει ότι όταν έχουμε υψηλό confussion χρησιμοποιούμε λάθος τις αντωνυμίες.
Casual cohesion από student	Flow	.598 > 0	Εδώ ο υψηλότερος θετικός συντελεστής δείχνει ότι όταν υπάρχει ροή στις απαντήσεις συνήθως έχουμε και casual σύνταξη κειμένου.

Noun overlap adjacent sentence από student	Frustration	$-.536 < 0$	Εδώ έχουμε τον χαμηλότερο αρνητικό συντελεστή που δείχνει ότι όταν έχουμε θυμό, έχουμε και ασύνδετες προτάσεις χωρίς νόημα.
--	-------------	-------------	---

22. Πως προκύπτει το 26% (...It appears that the cohesion predictors explained 26% of the variance averaged across the four affective states.)

Το 26% αναφέρεται στο adj. R squared το οποίο το αναλύσαμε παραπάνω. Πιο συγκεκριμένα έχουμε μέσω του πίνακα 1 ότι η τιμή του εξαρτάται από 4 συναισθηματικές καταστάσεις χρησιμοποιώντας τους παράγοντες πρόβλεψης που μετρούν τη συνεκτικότητα του κειμένου (Incident of negations από τον Student, Pronoun referential, cohesion από tutor, Casual cohesion από student, Noun overlap adjacent sentence από student) παίρνοντας τον μ.ο. των 4 μετρήσεων από τον πίνακα 1 σειρά 4, θα έχω $(0.246+0.223+0.33+0.258)/4 = 0.26425$ περίπου δηλαδή ίσο με 26%.

23. Αναφέρουν «...This is consistent with a medium to large effect for a statistical power of .8 ...» Εξηγήστε. Η απάντησή σας θα πρέπει να τεκμηριωθεί με αναφορά σε ιστοσελίδα που δίνει την ερμηνεία του όρου statistical power.

Σύμφωνα με το <https://machinelearningmastery.com/statistical-power-and-power-analysis-in-python/> και το κείμενο, έχουμε ότι:

Εδώ οι συγγραφείς κάνουν στατιστικό έλεγχο υποθέσεων με μηδενική υπόθεση (H_0) ότι δεν υπάρχει σχέση μεταξύ των συναισθημάτων των μαθητών και των μέτρων συνοχής του κειμένου μεταξύ του μαθητή και του tutor, και H_1 υπόθεση ότι υπάρχει σχέση ανάμεσα στα 2 και εξετάζει την στατιστική ισχύ του ελέγχου.

Η στατιστική ισχύς ή η ισχύς μιας δοκιμής υπόθεσης είναι η πιθανότητα ότι το τεστ απορρίπτει σωστά τη μηδενική υπόθεση. Δηλαδή η πιθανότητα ενός αληθινού θετικού αποτελέσματος. Είναι χρήσιμο μόνο όταν η μηδενική υπόθεση απορρίπτεται. Όσο υψηλότερη είναι η στατιστική ισχύς για ένα δεδομένο πείραμα, τόσο μικρότερη είναι η πιθανότητα να γίνει σφάλμα τύπου II (ψευδώς αρνητικό). Εδώ η ισχύς του ελέγχου είναι 0.8, που σημαίνει ότι αν η εναλλακτική υπόθεση είναι στη πραγματικότητα αληθής, σε 100 επαναλήψεις του ελέγχου των υποθέσεων, 80 φορές θα απορριφτεί η αρχική υπόθεση και θα γίνει αποδεκτή η εναλλακτική υπόθεση.

Το αν η αποδοχή της υπόθεσης έχει όντως νόημα επηρεάζεται από πολλούς παράγοντες, όπως το effect size. Μπορούμε να εκτιμήσουμε το effect size μέσω του R^2 και πιο συγκεκριμένα με το Cohens $f^2 = R^2/(1-R^2)$. Ο Cohen προτείνει ότι τιμές 0.02, 0.15, 0.35 το f^2 εκφράζουν μικρά, μεσαία και μεγάλα effect sizes αντίστοιχα. Αρα από το 26% που βρήκαμε πριν, βάζοντας στον τύπο θα προκύψει $f^2 \geq 0.35$ (γιατί έχουμε κάνει το adj R sq), τιμή που αντιστοιχεί σε μεγάλο effect size και δικαιολογεί το συμπέρασμα των συγγραφέων.

General Discussion

24. Οι συγγραφείς θέτουν το θέμα της γενίκευσης των μοντέλων πρόβλεψης όσον αφορά στο αντικείμενο στο οποίο εφαρμόστηκε (ερωτήσεις Πληροφορικής). Πως έλεγξαν την εγκυρότητα των αποτελεσμάτων τους;

Οι συγγραφείς προσπάθησαν να κάνουν έλεγχο για το αν οι 3 τομείς πληροφορικής επηρεάζουν τα μοντέλα παλινδρόμησης. Έτσι, δημιούργησαν ένα schema ελέγχου 2 βημάτων για να τσεκάρουν την στατιστική σημαντικότητα των παλινδρομήσεων:

- i. Δημιούργησαν dummy coded μεταβλητές για τις 3 κατηγορίες τομέων και χρησιμοποιήθηκαν ως παράγοντες της παλινδρόμησης. Όλα τα μοντέλα του βήματος ήταν στατιστικά μη σημαντικά
- ii. Χρησιμοποιήθηκαν ως παράγοντες αποκλειστικά οι συναισθηματικοί δείκτες του πίνακα 1. Όλα τα μοντέλα ήταν στατιστικά σημαντικά.

Έτσι φαίνεται ότι τα αποτελέσματα είναι ανεξάρτητα από τον τομέα, οπότε λογικά θα μπορεί να γίνει εφαρμογή και σε άλλους τομείς, χωρίς να έχει γίνει σίγουρη επικύρωση.

25. Ποιο μελλοντικό στόχο θέτουν οι συγγραφείς;

Να γίνει η παραπάνω διαδικασία σε real time, μέσω dialog windows του tutor κατά την διάρκεια της λύσης.

26. Διατυπώστε το σκοπό της εργασίας (μέγιστο πλήθος λέξεων 100)

Στην εργασία αυτή διερευνάται αν μπορούν να γίνει πρόβλεψη των συναισθημάτων των μαθητών, όπως η boredom, το flow, το confusion και η απογοήτευση μέσω μελέτης των διακυμάνσεων της συνεκτικότητας που παρατηρούνται σε διαλόγους που προκύπτουν από την αλληλεπίδραση μεταξύ του tutor και του εκπαιδευόμενου με τη βοήθεια του λογισμικού AutoTutor. Για την μελέτη των συναισθημάτων δεν γίνεται real time ανάλυση, αλλά ανάλυση με την βοήθεια 4 κριτών (ο εαυτός του, οι συμμαθητές και 2 ειδικοί) και η στατιστική ανάλυση έγινε με την βοήθεια του coh metrix program. Τέλος, μελετήθηκαν ποια μέτρα συνοχής ήταν τα πιο αποτελεσματικά για την πρόβλεψη των συναισθηματικών καταστάσεων, όπου χρησιμοποιήθηκαν μοντέλα πολλαπλής παλινδρόμησης, αλλά και αν η διαδικασία μπορεί να γενικευτεί περαιτέρω η διαδικασία.