

APLIKASI CROSS VALIDATION PADA MODEL SKILL SISWA

Wahyu Hartono

Universitas Swadaya Gunung Djati, Cirebon, Indonesia
wahyuhartono.id@gmail.com

Abstract

One of the activities in the educational test is making a diagnosis to determine whether or not a person's skills are present. This study specifically aims to design student skill models in basic mathematics courses and perform validation using a leave-one-out cross validation to select an accurate model. The diagnostic test questions used in this study ranged from moderate to difficult. The findings of this study indicate that the method of fixed test questions in order of questions from easy to difficult is better than the method of design of the initial fixed test questions..

Keywords: Bayes Network, Cross Validation, Student Skill Model, Diagnostic test

1. Pendahuluan

Ide inti dari penguasaan belajar adalah bahwa semua siswa dapat mencapai keahlian dalam materi belajar (domain) jika dua kondisi berikut terpenuhi: (1) pengetahuan domain secara tepat dianalisis ke dalam hierarki keterampilan komponen dan (2) pengalaman belajar disusun untuk memastikan bahwa siswa menguasai keterampilan prasyarat sebelum mempelajari keterampilan tingkat yang lebih tinggi dalam hierarki, Corbett dan Anderson (1995).

Selama puluhan tahun, para peneliti telah mengembangkan berbagai metode untuk memodelkan siswa. Dua pendekatan yang paling terkenal adalah *Bayesian Knowledge Trace* (BKT) yang menggunakan *dynamic Bayesian Network* untuk memodelkan pengetahuan atau *skill* siswa, dan *Performance*

Factor Analysis (PFA), yang menggunakan regresi logistik untuk memprediksi *student performance*, Wang, Y dan Beck, J (2018).

Salah satu kegiatan dalam tes pendidikan adalah melakukan diagnosa untuk mengetahui ada atau tidak adanya *skill* yang dimiliki seseorang. Skenario yang mungkin adalah lembaga pendidikan menyediakan berbagai perkuliahan dan melaksanakan tes diagnosa kepada calon mahasiswa untuk mengetahui materi prasyarat apa saja yang belum dimiliki kemudian berusaha menguasainya agar bisa mengikuti perkuliahan tersebut.

Menurut Vomlel (2004), dalam merancang suatu tes, dibentuk sebuah himpunan yang terdiri dari keahlian (*skill*), kemampuan (*ability*), miskonsepsi, dan seterusnya, serta sebuah himpunan yang berisi bank soal, tugas, dan seterusnya. Misalkan $S = \{S_1, S_2, \dots, S_k\}$ menotasikan *skill*, *ability*, miskonsepsi, dan lain-lain, serta $\chi = \{X_1, X_2, \dots, X_m\}$ menotasikan bank soal, tugas, dan lain-lain. Untuk mempermudah selanjutnya S disebut *skills* dan χ disebut pertanyaan. Perancang tes harus menentukan *skill* mana yang secara langsung berhubungan dengan setiap pertanyaan. Hubungan tersebut seringkali bersifat probabilitas, khususnya pada soal tes berbentuk pilihan ganda.

Salah satu pendekatan dalam merancang suatu tes adalah menyusun serangkaian pertanyaan yang mengakomodasi seluruh *skills* yang akan diuji. Pendekatan tersebut sering disebut sebagai *fixed test*. Pendekatan lainnya adalah menyusun soal tes yang optimal untuk masing-masing peserta tes. Setelah suatu pertanyaan dijawab oleh peserta tes, sistem akan memilih pertanyaan selanjutnya berdasarkan jawaban dari pertanyaan sebelumnya. Karena pendekatan tersebut menggunakan komputer dalam penyajian tes nya, maka sering disebut *Computer Adaptive Testing* (CAT), Mislevy (2000). Tes yang secara otomatis menyajikan pertanyaan

berdasarkan jawaban dari pertanyaan sebelumnya disebut sebagai *adaptive tests*.

Jaringan Bayes (*Bayesian Networks*) adalah model yang populer dari keluarga model berbentuk grafik, Lauritzen (1996). Beberapa peneliti telah mengaplikasikan Jaringan Bayes untuk tes pendidikan dan pelatihan. Millan et al. (2002) menggunakan Jaringan Bayes untuk memodelkan siswa pada *computerized adaptive tests*. Mislevy et al. (1999) meneliti bagaimana parameter numerik dari model probabilitas dapat ditaksir. Almond et al. (2001) meneliti model untuk tabel-tabel probabilitas bersyarat pada penilaian pendidikan. Dalam karya ilmiah yang ditulis oleh Conati et al. (1997) aplikasi menarik dari Jaringan Bayes untuk memodelkan siswa dalam melatih kemampuan pemecahan masalah telah disajikan. Penelitian ini secara khusus bertujuan untuk merancang model skill siswa pada mata kuliah matematika dasar dan melakukan validasi menggunakan *cross validation leave-one-out* untuk memilih model yang akurat.

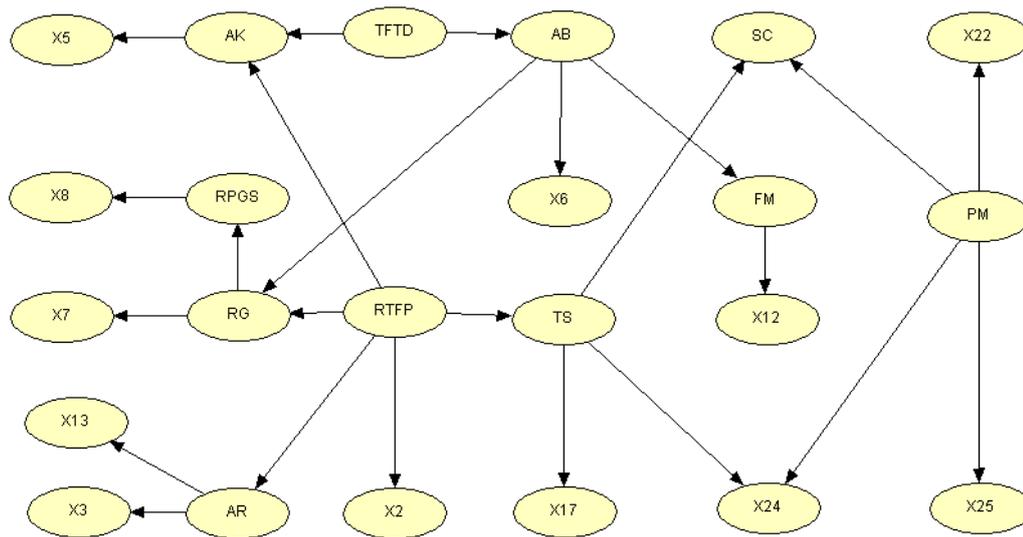
2. Metode Penelitian

Penelitian ini merupakan penelitian eksperimen. Data-data hasil tes dan Model skill siswa terkait materi Kalkulus Diferensial yang digunakan dalam penelitian ini berasal dari penelitian (Hartono dan Tonah, 2017). Validasi model struktur belajar menggunakan *Cross Validation Leave-One-Out* karena model yang terbangun sangat kompleks dan labelnya sulit ditentukan. Dari hasil validasi selanjutnya akan dipilih model yang paling akurat. Model tersebut akan digunakan untuk memprediksi kemampuan matematika dasar siswa. Kesimpulan yang akan diperoleh adalah informasi *prediction error* dari setiap model skill siswa yang telah dibuat.

3. Hasil dan Pembahasan

A. Hasil Pengolahan Data

Model skill siswa terkait materi Kalkulus Diferensial yang digunakan dalam penelitian ini berasal dari penelitian (Hartono dan Tonah, 2017). Model tersebut dapat dilihat pada Gambar 4.1 berikut.



Gambar 4.1 Model siswa yang menggambarkan hubungan antara skill dan soal

Pada Gambar 4.1 Variabel X melambangkan nomor soal, sedangkan variabel lainnya merupakan singkatan dari skill/kemampuan matematis yang akan diprediksi dimiliki oleh mahasiswa atau tidak. Tanda panah keluar berarti bahwa skill tersebut dibutuhkan oleh skill/soal diujung tanda panah. Tabel 4.1 berikut merupakan kepanjangan dari singkatan pada Gambar 4.1.

Tabel 4.1 Kode Skill Siswa

No.	Skill	Deskripsi
1	RTFP	Rumus Turunan Fungsi Polinom
2	AR	Aturan Rantai
3	AK	Aturan Perkalian

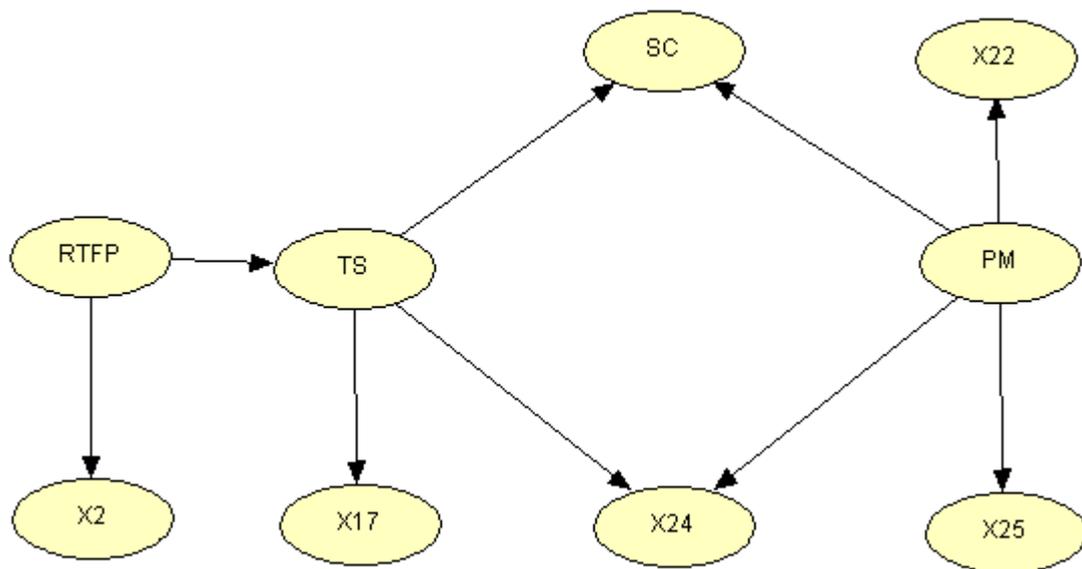
No.	Skill	Deskripsi
4	AB	Aturan Pembagian
5	TFTD	Turunan Fungsi Trigonometri Dasar
6	RG	Rumus Gradien
7	RPGS	Rumus Persamaan Garis Singgung
8	FM	Fungsi Monoton
9	TS	Titik Stasioner
10	PM	Pemodelan Matematika
11	SC	Soal Cerita

Soal *Fixed Test* yang sesuai dengan model pada Gambar 4.1 diberikan kepada 150 mahasiswa program studi pendidikan matematika Universitas Swadaya Gunung Djati. Jawaban siswa yang benar semua atau salah semua dihilangkan untuk menghilangkan bias perhitungan dari sampel. Sampel tersebut akan digunakan untuk melakukan *Leave One Out Cross Validation* dengan metode Klasifikasi Naïve Bayes. Tingkat kesukaran butir dari soal ujicoba tersaji pada Tabel 4.2 berikut.

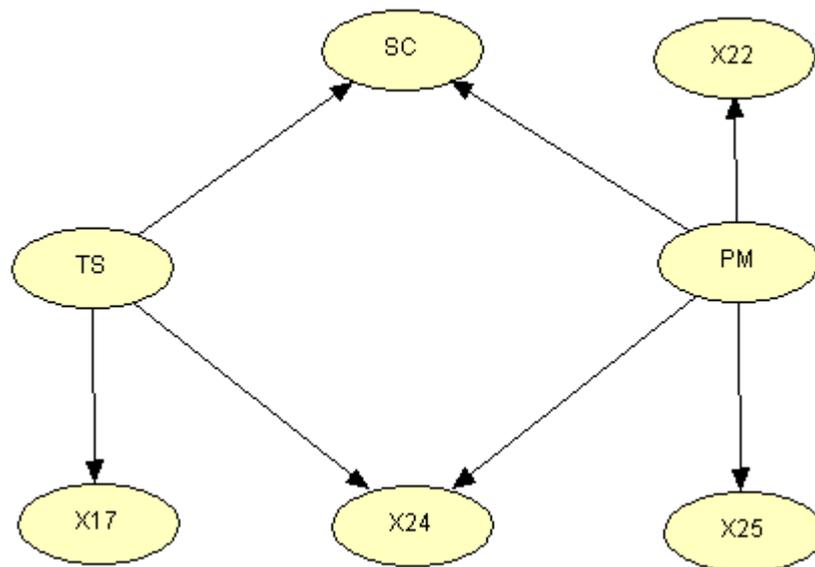
Tabel 4.2 Tingkat Kesukaran Butir

No.	Butir Soal	Tingkat Kesukaran	Kategori Soal	Urutan Mudah ke Sukar
1	X2	0,360	Sedang	1
2	X3	0,213	Sukar	2
3	X5	0,087	Sukar	9
4	X6	0,213	Sukar	3
5	X7	0,080	Sukar	11
6	X8	0,047	Sukar	12
7	X12	0,020	Sukar	14
8	X13	0,040	Sukar	13

9	X14	0,120	Sukar	5
10	X17	0,080	Sukar	10
11	X21	0,107	Sukar	6
12	X22	0,100	Sukar	7
13	X24	0,093	Sukar	8
14	X25	0,133	Sukar	4



Gambar 4.2 Model siswa yang menggambarkan hubungan antara skill dan soal yang berkontribusi dengan node SC.



Gambar 4.3 Model siswa yang menggambarkan hubungan antara skill dan soal yang langsung berkontribusi dengan node SC.

Untuk mempermudah pembahasan, selanjutnya Gambar 4.1 disebut sebagai model 1, Gambar 4.2 sebagai model 2, dan Gambar 4.3 sebagai model 3. Peneliti menggunakan Software R untuk mengimplementasikan *Leave One Out Cross Validation* (LOOCV) dengan metode klasifikasi Naive Bayes. Program R untuk LOOCV dengan metode Naive Bayes dapat dilihat pada Lampiran 1. Tabel 4.3 berikut adalah hasil implementasikan *LOOCV* pada metode klasifikasi Naive Bayes untuk model 1, model 2, dan model 3.

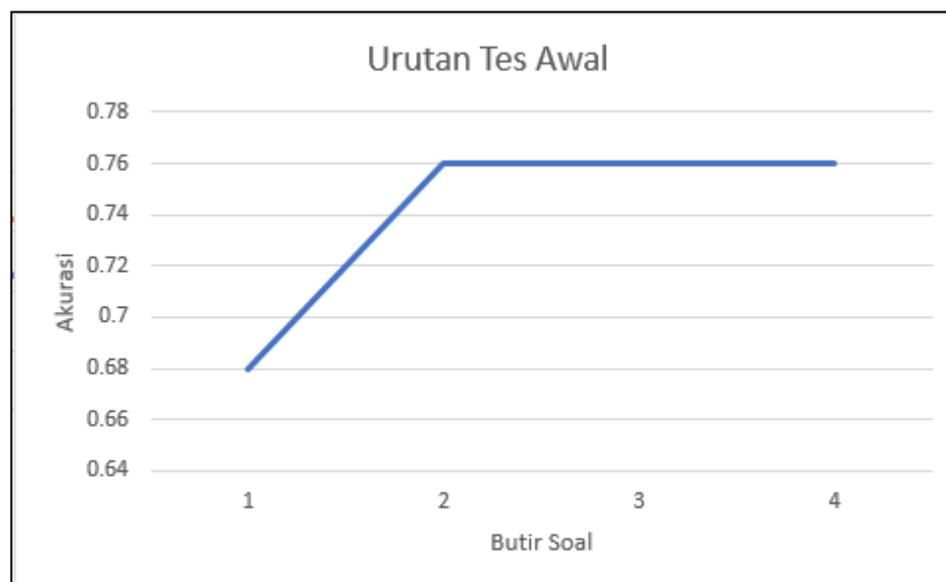
Tabel 4.3 Akurasi Model

Model	Akurasi
Model 1	0,465
Model 2	0.765
Model 3	0,800

Berdasarkan hasil pada Tabel 4.3 dipilih Model 3 karena tingkat akurasi yang paling tinggi. Selanjutnya akan dibandingkan desain soal fixed test dengan urutan soal awal dan urutan soal dari mudah ke sukar yang sesuai dengan Model 3.

B. Desain Soal Fixed Test Awal

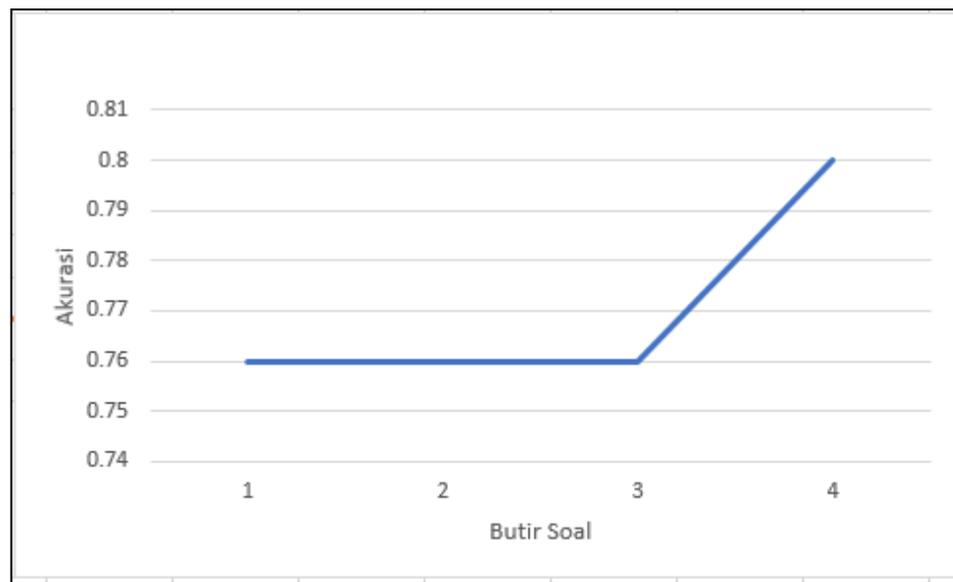
Urutan butir soal dari Model 3 yang sesuai tes awal adalah X17, X22, X24, X25. Butir-butir soal tersebut akan digunakan untuk mengestimasi apakah seorang siswa memiliki kemampuan untuk menyelesaikan soal cerita (SC) terkait materi kalkulus diferensial. Dalam terminologi klasifikasi dengan metode Naïve Bayes, butir-butir X17, X22, X24, X25 disebut Feature, sedangkan SC disebut variabel target. Setelah mengimplementasikan LOOCV dengan metode Naïve Bayes, diperoleh tingkat akurasi untuk setiap jumlah butir seperti pada Gambar 4.4 berikut.



Gambar 4.4 Prediksi Skill SC Menggunakan Tes Awal

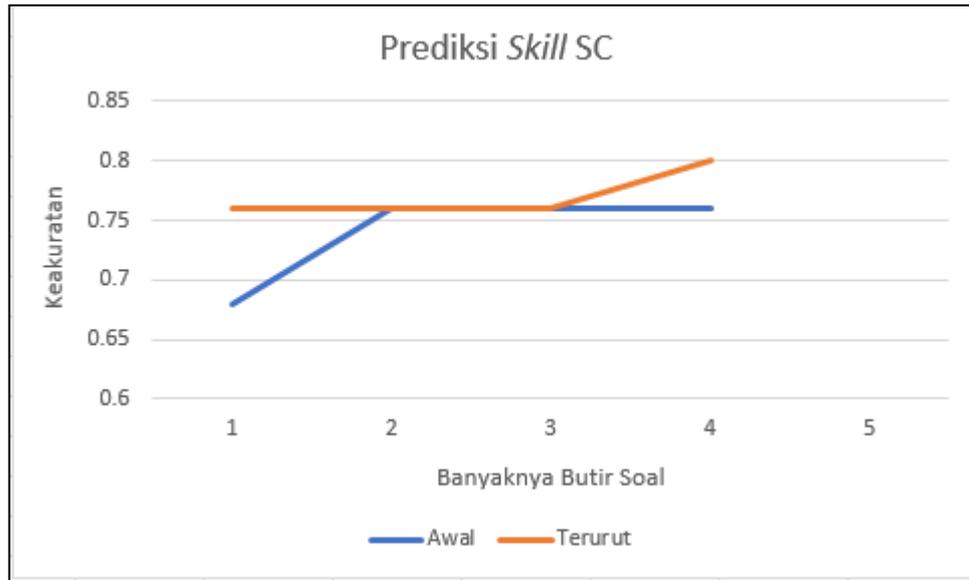
C. Desain Soal Fixed Test dari Mudah ke Sukar

Urutan butir soal dari Model 3 yang urutan butir tes nya dari mudah ke sukar adalah X25, X22, X24, X17. Butir-butir soal tersebut akan digunakan untuk mengestimasi kemampuan siswa untuk menyelesaikan soal cerita (SC) terkait materi kalkulus diferensial. Setelah mengimplementasikan LOOCV dengan metode Naïve Bayes, diperoleh tingkat akurasi untuk setiap jumlah butir seperti pada Gambar 4.5 berikut.



Gambar 4.5 Prediksi Skill SC dengan urutan butir dari mudah ke sukar

Jika semua prediksi skill dari desain fixed test awal dan terurut dari mudah ke sukar maka akan diperoleh hasil seperti pada Gambar 4.6 berikut.



Gambar 4.6 Prediksi Skill SC dari tes awal dan tes terurut

Pada Gambar 4.6 terlihat bahwa kinerja fixed test yang butir soalnya terurut dari mudah ke sukar lebih baik daripada desain soal *fixed test* awal.

4. Kesimpulan

Temuan penelitian ini menunjukkan bahwa metode *fixed test* yang urutan soalnya dari mudah ke sukar lebih baik daripada metode desain soal *fixed test* awal. Soal tes diagnosa yang digunakan dalam penelitian ini memiliki tingkat kesukaran dari sedang ke sukar. Saat mendesain tes diagnosa sebaiknya diperhatikan sebaran soal agar soal diagnosa memiliki sebaran dari yang mudah ke yang sukar. Banyaknya soal ditambah agar akurasi prediksinya dapat mendekati angka 1.

DAFTAR PUSTAKA

- Almond, Russell G dan Mislevy, Robert J. 1999. Graphical models and computerized adaptive testing. *Applied Psychological Measurement*, 23(3):223–237.
- Corbett, A., Anderson, J. 1995. Knowledge Tracing: Modeling the Acquisition of Procedural Knowledge. *User Modeling and User-Adapted Interaction* 4, 253–278.

- Conati, Cristina et al. 1997. On-line student modeling for coached problem solving using Bayesian networks. In Anthony Jameson, Cecile Paris, and Carlo Tasso, editors, Proc. of the Sixth Int. Conf. on User Modeling (UM97), Chia Laguna, Sardinia, Italy, pages 231–242. Springer Verlag.
- Millan, Eva dan Pérez-de-la-Cruz, José Luis. 2002. A Bayesian diagnostic algorithm for student modeling and its evaluation. *User modeling and User-Adapted Interaction*, 12(2–3):281–330.
- Mislevy, Robert J et al. 2000. *Computerized Adaptive Testing: A Primer*. Mahwah, N.J., Lawrence Erlbaum Associates, second edition.
- Hadi, S. 2013. *Pengembangan Computerized Adaptive Test Berbasis Web*. Yogyakarta: Aswaja Pressindo
- Vomlel, Jiri . 2004. Bayesian Networks in Educational Testing. *International Journal of Uncertainty, Fuzziness and Knowledge Based Systems*, Vol. 12, Supplementary Issue 1, 2004, pp. 83-100. [A draft version](#).
- Ben-Bassat, Moshe. 1978. Myopic policies in sequential classification. *IEEE Transactions on Computers*, 27(2): 170–174.
- Almond et al. 2001. Models for conditional probability tables in educational assessment. In Proc. of the 2001 Conference on AI and Statistics. Society for AI and Statistics.
- Lauritzen, Steffen L. 1996. *Graphical Models*. Clarendon Press, Oxford.
- Plajner, M dan Vomlel, J (2016). Student Skill Models in Adaptive Testing. *JMLR: Workshop and Conference Proceedings vol 52, 403-414*.
- Wang, Y., Heffernan, N.T. 2012. The Student Skill Model. In: Cerri, S.A., Clancey, W.J., Papadourakis, G., Panourgia, K. (eds.) ITS 2012. LNCS, vol. 7315, pp. 399–404. Springer, Heidelberg.
- Wang Y., Beck J. (2013) Class vs. Student in a Bayesian Network Student Model. In: Lane H.C., Yacef K., Mostow J., Pavlik P. (eds) *Artificial Intelligence in Education. AIED 2013. Lecture Notes in Computer Science*, vol 7926. Springer, Berlin, Heidelberg