Better Poster and ÚFAL Templates

Tom Kocmi, Tomáš Musil

September 17, 2019
Outline

1. Redesigning poster
2. Better poster design
3. What works and what not so much
4. ÚFAL templates
Preparing Poster
Preparing Poster

- pack the most information as possible
- designing poster understandable after 10 or more minutes
- doing the poster last minute
Non-Cognitive Predictors of Student Success: A Predictive Validity Comparison Between Domestic and International Students
Jacob Smith, Dr. Thea Schofield, Dr. Antonio Ibarra, Stephen Choi, Bean Mullins, Dr. Emily Williams
Michigan State University

Abstract

Given increasing interest in utilizing non-cognitive predictors in the college admissions process and rising enrollment of international students, research is warranted to compare the predictive validity of these measures across domestic and international students. Results indicate some predictive validity differences do exist, and an implication for this differential validity, as well as a moderator of these relationships, are tested.

Background

Though cognitive predictors of student success (e.g., ACT, HSPTs) remain popular, there is a growing interest in non-cognitive predictors of student success (e.g., situational judgment, adaptability), and these have been found to predict student performance (Covello et al., 2004; Kamy et al., 2005). From 2008 to 2018 academic years, the number of international students studying in the U.S. increased annually. In 2018, 2.5% of students internationally with over one million enrolled (Institute of International Education, 2016). Previous work by Prasad and colleagues (2018) found mean differences in non-cognitive measures across Chinese and Cajun American students, along with differential validity for a Perseverance non-cognitive measure. The current research is an extension of Prasad et al., 2018, exploring differential validity in two large samples of students, testing an explanation for these differences in validity, and noting a potential interaction of these relationships between non-cognitive predictors and GEDs.

Research Question & Hypotheses

Research Question 1: Will non-cognitive measures display differential validity between domestic and international students?
H1: Differential validity will be accounted for by English proficiency.
H2: Non-cognitive predictors may be more important for individuals from a more culturally distant country, as adjustment may be more difficult accomplishing.
H3: Non-cognitive measures will exhibit greater validity for international students from more culturally distant countries.

Method (cont.)

Table 2. Moderated Regression Results for Non-Cognitive Predictors of Student Success

<table>
<thead>
<tr>
<th>Predictor</th>
<th>Overall Sample</th>
<th>Domestic Sample</th>
<th>International Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>GED</td>
<td>0.25</td>
<td>0.23</td>
<td>0.27</td>
</tr>
<tr>
<td>SAT</td>
<td>0.34</td>
<td>0.32</td>
<td>0.36</td>
</tr>
<tr>
<td>ACT</td>
<td>0.30</td>
<td>0.28</td>
<td>0.32</td>
</tr>
<tr>
<td>GPA</td>
<td>0.22</td>
<td>0.20</td>
<td>0.24</td>
</tr>
<tr>
<td>Age</td>
<td>0.18</td>
<td>0.16</td>
<td>0.20</td>
</tr>
<tr>
<td>Gender</td>
<td>0.10</td>
<td>0.08</td>
<td>0.12</td>
</tr>
<tr>
<td>Race</td>
<td>0.05</td>
<td>0.03</td>
<td>0.07</td>
</tr>
</tbody>
</table>

**Bold numbers indicate significant relationships (p < .05)**

Correlations between non-cognitive predictor scores and 1st semester GPA (Table 1) indicate stronger relationships for international students on six of eight measures. Regression results (Table 2) indicate consistent differential validity for the international students for SAT, Continuous Learning, Social Responsability, and Perseverance. Including TOEFL scores in regression analysis, available for a subset of 81 individuals from Sample 1, did not substantively alter standardized regression weights (AB = -.01 to .03). (Results not shown).

Multiple regression was utilized to test cultural distance via GLOBE moderated validity for non-cognitive predictors utilizing subset of 21 international students from Sample 1 from 11 countries. Results indicate cultural distance did not significantly moderate validity (p > .05). (Results not shown).

Contrasts between GLOBE cultural distances and perceived cultural distance **...**

Discussion

Results indicate consistent differential validity for some non-cognitive measures for international students, specifically for SAT, Continuous Learning, Social Responsability, and Perseverance. Differential validity for international students does not seem to be the results of functioning as a proxy for English language ability. Cultural distance does not seem to moderate validity of non-cognitive measures.

Implications

Non-cognitive abilities may be useful in predicting international student performance, but differential validity may be an issue.
Negates non-significant relationship between cultural distance via GLOBE scores and perceived cultural distance warrants caution in generalizing country level scores to individuals.
More research is warranted to explain differential validity for international students.

Acknowledgments

I would like to thank Disney Macias for assistance in data collection, as well as Frank Zheng and Bob Duffin for advice regarding data analysis.

“Perfection is not when you have nothing to add, it is when you have nothing to take away.” — Beethoven
Attendee Point of View
Goals of Attendee

- Planning to learn the most
- Planning to talk in depth (15 minutes) with 2-3 people
- Restricted time (90 minutes)
- Too many posters (40-50 posters per session of ACL 2019)
- Leaving 2 minutes per poster > Need to skim through most posters
Perfection is not when you have nothing to add, it is when you have nothing to take away.
1. Maximize first insight
2. Keep only the good stuff
3. Easy and fast to create
Better Poster Concept
We found consistent differential validity for some non-cognitive measures for predicting international student GPA, specifically with SJT, Continuous Learning, Social Responsibility, and Perseverance.
For international students, perseverance and a sense of social responsibility are extra important for predicting first-year GPA.
For international students, perseverance and a sense of social responsibility are extra important for predicting first-year GPA.
Non-Cognitive Predictors of Student Success: A Predictive Validity Comparison Between Domestic and International Students

• Jacob Smith, Dr. Theo Schofield, Dr. Antonio Barr, Jane Choi, Bruce Mullen, Dr. Emily Williams

INTRO
• Increasing interest in utilizing non-cognitive predictors in the college admissions process
• Rising enrollment of international students

METHODS
• We compare the predictive validity of these measures across domestic and international students.
• Results indicate some predictive validity differences do exist and can explain for this differential validity, as well as a moderator of these relationships, are tested.

RESULTS
• Consistent differential validity for some non-cognitive measures for international students, specifically for Q1, Continuous Learning, Social Responsibility, and Perseverance.
• Differential validity for international students based on the results of the scale, indicating the necessity of adjusting as a proxy for English-language ability.
• Cultural distance does not seem to moderate validity of non-cognitive

DISCUSSION
• Non-cognitive abilities may be useful in predicting international student performance, but differential validity may be an issue.
• Negative, non-significant relationship between cultural distance via GLOBAL scores and perceived cultural distance warrants caution in generalizing country-level scores to individuals.
• More research is warranted for English-differential validity for international students.

For international students, perseverance and a sense of social responsibility are extra important for predicting first-year GPA.
Non-Cognitive Predictors of Student Success:
A Predictive Validity Comparison Between Domestic and International Students

INTRO
- Increasing interest in utilizing non-cognitive predictors in the college admission process
- Rising enrollment of international students

METHODS
- We compare the predictive validity of these measures across domestic and international students.
- Results indicate some predictive validity differences do exist and an explanation for this differential validity, as well as a moderator of these relationships, are tested.

RESULTS
- Consider differential validity for some non-cognitive measures for international students, specifically for EFL, Extraversion, Openness, and Persistence.
- Differential validity for international students may be due to the impact of English language proficiency as a proxy for English language ability.
- Culture and language do not seem to moderate validity of non-cognitive predictors in international students.

DISCUSSION
- Non-cognitive abilities may be useful in predicting international student performance, but differential validity may be an issue.
- Negative, non-significant relationship between cultural distance and GLC scores and prosocial/cultural distance war with caution in generalizing country-level scores to individuals.
- More research is warranted to explain differential validity for international students.

For international students, perseverance and a sense of social responsibility are extra important for predicting first-year GPA.
What Really Works in NLP
What Really Works in NLP

Non-Cognitive Predictors of Student Success: A Predictive Validity Comparison Between Domestic and International Students

- Jacob Smith, Dr. Theo Schmid
- Dr. Antonio Burt, Jane Cho, Ben Mullins
- Dr. Emily Williams

INTRO
- Increasing interest in utilizing non-cognitive predictors in the college admissions process
- Rising enrollment of international students

METHODS
- We compare the predictive validity of these measures across domestic and international students.
- Results indicate some predictive validity differences do exist and an explanation for this differential validity, as well as a moderator of those relationships, are tested.

RESULTS
- Considered differential validity for some non-cognitive measures for international students, particularly for GED, Continuous Learning, and Social Integration.
- Differential validity for international students is found using a composite of standardized test scores, which is used as a proxy for English language ability.
- Cultural differences do not seem to moderate validity of non-cognitive measures.

DISCUSSION
- Non-cognitive abilities may be useful in predicting international student performance, but differential validity may be an issue.
- Negative, non-significant relationship between cultural distance and some cultural distance welfare indices in generating country-level scores to individuals.
- More research is warranted to explore differential validity for international students.
Real Examples
Towards Integration of Statistical Hypothesis Tests into Deep Neural Networks

Dr. Raul Agustadimesco
Zurich University of Applied Sciences

Real Examples

Introduction
We propose a new deep architecture which works in tandem with a statistical test procedure for training a text classifier jointly on texts and their labeled descriptions. The model leverages the use of label descriptors in addition to the input text to improve text classification performance.

Intuition
The intuition is to help the model to concentrate on more informative words rather than more frequent ones hence enhancing the model performance.

Conclusion
Using statistical hypothesis testing methods, we extract the most informative words for each class as the class descriptors. We also propose an architecture for actively attending over class descriptors hence improving the state-of-the-art in text classification. Our method is entirely data-driven, has no dependency on other sources of information than the training data, and is adaptable to different classification problems by providing appropriate training data without major hyperparameter tuning.

Jointly Training Deep Neural Classifiers on Text and $\chi^2$ Labels using $\chi^2$ Hypothesis test

Better Poster and ÚF AL Templates

Preparing Poster
Attendee Point of View
Perfection is not when you have nothing to add, it is when you have nothing to take away.

Beamer Template for ÚF AL

12/21
Unsupervised word segmentation models perform above chance & stably across languages → They could be potential strategies employed by infants.

Is word segmentation child’s play in all languages?

**Georgia Lavakis**, Steven Moran, Daniel Beli, Sabine Veale, Alex Crista

**RESULTS**
- DBS performed above baselines in all languages. The others performed above baselines for nearly all languages.
- Spearman corr showed a similar rank ordering of model performance across languages. Only Innsitut and Russian differed.
- Better differences in performance across models (Table 2) than across languages (Table 2).

**METHODS**
- WordSeg package
  - 2 baseline (cut anywhere at p<0.15)
  - DBS: diphone prob, optimal choice
  - TTS: diphone prob (backwards, FW, M)
  - AG: lexical & generative
  - PUDLE: test & memory-based
  - AuxDiv database

**DISCUSSION**
- Many proposed models perform above chance & stably across languages.
- Infants used similar strategies, they could get a head start in segmenting word-like units regardless of what their native language is.

For cross-lingual word embeddings, a little preprocessing can drastically increase word translation accuracy.

Are Girls Neko or Shōjo? Cross-Lingual Alignment of Non-Isomorphic Embeddings with Iterative Normalization

Mozhi Zhang, Keyulu Xu, Kenichi Kawarabayashi, Stefanie Jegelka, Jordan Boyd-Graber

**Background**
- How to train bilingual word embeddings
  * Train two monolingual embeddings
  * Align with an orthogonal mapping
    - Definition: $W W^T = I$
    - Preserve dot-products: $x^T y = (Wx)^T(Wy)$

**Iterative Normalization**
- Preprocess monolingual embeddings before learning the alignment.
- For $k = 1, 2, \ldots$:
  - Length normalization:
    
  - Mean centering:
    
**Word translation accuracy**

Before | After
---|---
EN-JA | 1.7 | 44.2
EN-ZH | 32.5 | 44.2
EN-HI | 33.3 | 36.7
EN-TR | 44.9 | 48.7
EN-DA | 54.0 | 58.8

*Results on MUSE*
- Improve on 29 target languages
- Helpful for both orthogonal and non-orthogonal methods (such as RCoSL)

**Why does it work?**
- Two conditions that help orthogonal mapping:
  1. Every vector has the same length.
  2. Each language’s mean has the same length.
- Condition 1 unifies objectives: dot-product for training monolingual embedding, cosine similarity for evaluating monolingual embedding, and Euclidean distance for cross-lingual alignment.
- Condition 2 is a necessary condition for orthogonal alignment.
- Iterative Normalization uses alternating projection to provably satisfy both.
We developed a flexible, terminal-based annotation tool that takes seconds to install and is easy to extend.

SLATE: A Super-Lightweight Annotation Tool for Experts

Jonathan K. Kummerfeld

Overview
A new annotation tool designed for users with a terminal-based workflow, with:
- Thoral installation
- Simple flat file format
- Annotation at different scales (characters, tokens, lines, document)
- Annotation on different types (free text, labels, links)
- Customizable keybindings
- Code that is easy to modify to add features
- Unicode support (only used for English projects so far)

Install and use:
Option 1: pip install slate-clp
Option 2: git clone https://github.com/jkummerfeld/slate

Labeling Cybercriminal Money Exchange
- Current labels selected:
- Green:ассив
- Blue:Labelled text, as indicated at the bottom
- Underline:
- Light/ dark blue: Annotation tags disagree / agree on
- Red: Other messages with disagreements

Example from "Tools for Automated Analysis of Cybercriminal Markets", Portnoy et al. [WWW 2017]

Adjudicating IRC Disagreements
- Green: Current message selected
- Light/ dark blue: Messages disagree/agree on
- Underline:
- Message being considered for linking to the group message

Example from "A Large-Scale Corpus for Conversation Disagreement", Kummerfeld et al. [ACL 2019]
Add Main Message in Plain English
Better Poster for ÚFAL
Main finding goes here, translated into plain English. Emphasize the important!

Features built from unlabeled data improve automatic evaluation of coherence

Introduction

We ask new possibilities of using large unlabeled data to improve quality of automatic evaluation of surface coherence in student essays. Particularly we propose two approaches to benefit from the large data in Cz language model, and already estimated features used to evaluate the features. In our experiments, we integrate these approaches with unlabeled data from the Czech National Corpus into the evaluator of surface coherence for Czech, the EVALD system, and test by performance on two decisive dataset written by native speakers (L1) as well as foreign learners of Czech (L2).

Labeled data

Unlabeled data

Contribution

• A new proposed approach of exploiting unlabeled data in combination with other new features improve the performance of EVALD.
• Increase in F-score is much higher for L1, probably due to higher variability of the unlabeled text in the L2 texts.
Better Poster for ÚFAL

Preparing Poster

Attendee Point of View

Perfection is not when you have nothing to add, it is when you have nothing to take away.
Better Poster for ÚFAL

Perfection is not when you have nothing to add, it is when you have nothing to take away.
Better Poster for ÚFAL

Preventing Poster Attendee Point of View Perfection is not when you have nothing to add, it is when you have nothing to take away.
Better Poster for ÚFAL

Perfection is not when you have nothing to add, it is when you have nothing to take away.
Summary

1. Add the main message in plain English
2. Do not be scared of negative space
3. ÚFAL consistency

https://wiki.ufal.ms.mff.cuni.cz/presentations