Linking Soft Skills and Labor Market Outcomes: Evidence from the NLSY97 and Machine Learning Techniques

Matt Gevercer, Zoe Krieger, Diego Saavedra, Benjamin Schneider, Benjamin Zwarg

June 10, 2022

Abstract

In developing countries like Ethiopia and South Africa more than fifty percent of youth are unemployed, many of whom are greatly disadvantaged in their respective labor markets by a lack of access to education. However, as soft skills are beginning to gain value more and more for workers across many industries, there is growing interest and increased investment in developing these non-technical skills. In this paper, we explore different measures of job fit and, using longitudinal labor data, we attempt to predict worker fit to different industries using psychometric data. Using multiple machine learning and econometrics methods we find compelling evidence that predictive matches can be made utilizing soft skills data.

Introduction

Understanding labor market outcomes is a complex and exciting problem in the field of economics. Much work has been done linking factors like education and hard skills (Deming 2017, etc.) to positive outcomes, leading to significant investment and focus on increasing access to education and skill-building opportunities. In particular, the World Bank is interested in improving labor market outcomes in developing countries. However, in countries like Ethiopia and South Africa more than fifty percent of youth are unemployed, often due to a lack of access to education. With limited resources to improve educational opportunities, the World Bank aims to promote disadvantaged entry-level job applicants by emphasizing non-cognitive soft skills in the hiring process.

While there is a growing body of literature linking general individual traits and skills to job outcomes, there is a significant lack of studies that aim to directly apply their findings to policies and programs in the field, specifically in the Human Resources setting. In addition to this there are very few studies or papers that solely focus on soft skills, and even fewer of those utilize machine learning techniques.

This paper therefore explores the link between soft skills and labor market outcomes and aims to accurately predict successful job matches with soft skills data. Specifically, we are using National Longitudinal Survey of the Youth 1997 (NLSY97) data on both occupation and personality traits to run our analyses. First we briefly describe previous literature on the topic and give an overview of the dataset we used. Then, we explain our methods in detail, as well as the different matching and predictive methods we employ. Finally, we discuss our findings and the implementation of our results for our specific World Bank setting.

Literature Review

While much research has been done in numerous fields linking education and technical skills to labor market outcomes, research on how soft skills determine labor market outcomes is considerably more narrow. Deming (2017) notably addresses this gap in the literature with his investigation of the changing importance of social skills over the past few decades. Specifically, through modeling and widely-used empirical methods, Deming finds social skills to be a significantly important predictor of both wages and full-time employment and also presents evidence that efficient sorting into jobs based on social skills leads to better labor market outcomes. However Deming (2017) relies on a very general definition of social skills, only measuring survey respondent's sociability, not considering other social skills or soft skills.

Many papers aim to build off of Deming (2017) and identify exactly which soft skills are linked with favorable market outcomes in a range of countries and environments (Cunningham and Villaseñor, 2014; Acosta et al., 2015). Most relevant to our research is Guerra and Cunningham (2014) which outlines current interdisciplinary findings linking social-emotional skills to labor market success, focusing on which skills employers value. The authors present the PRACTICE model which outlines which skills have been found most valuable to employers. The model lists eight main skills (Problem Solving, Resilience, Achievement Motivation, Control, Teamwork, Initiative, Confidence, Ethics), multiple sub-skills, and links each skill to a Big Five Personality Traits and GRIT. Guerra and Cunningham (2014) provide a strong basis of understanding the link between soft skills and labor market outcomes but the PRACTICE model must contend with other similar models such as the Soft Skills Inventory developed in Jardim et al. (2022) which uses data from Portugal and finds several different categories and different groupings of similar questions. It is clear that there is still a gap in understanding of how soft skills determine labor market outcomes.

Another study of interest, Kern et al. (2019) uses more modern techniques to link personality profiles and occupation. The authors rely on Big Five Personality Traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) scores derived from linguistic analysis of Twitter data and matched estimated personalities to nine occupations. Distinct digital fingerprints emerged from the data with greater alignment for similar occupations. For example they found that software programmers and chemistry teachers generally scored high on Openness but lower on Agreeableness and Conscientiousness. The authors then built a vocations map to cluster similar jobs together based on personality scores and similarities in job-task requirements. Finally they trained a machine-learning algorithm to predict good-fitting jobs based on personality scores. While many of the methods in Kern et al. (2019) are different from what we aim to do, the authors set precedence for the use of machine learning models to investigate the relationship between personality and labor market outcomes. We hope to extend this research to broader applications, filling the gap in the literature for evaluating soft skills and labor market outcomes in developing countries via machine learning.

Between the fields of econometrics and machine learning, there are a variety of techniques to relate psychometric attributes to job outcomes. Our particular application– namely, building a model that takes data from job applicants and assesses their “job fit” as a preliminary step for the hiring process– requires special consideration of each technique's advantages and drawbacks.

Varian (2014) notes that machine learning techniques are optimized for handling big data. While there is no consensus about what defines the term “big data,” Varian focuses on datasets with "several gigabytes of data or several million observations." Compared to these, our longitudinal datasets are relatively miniscule, especially after accounting for missing values and nonresponses.

Still, there is no requirement that machine learning models be trained with large datasets, and there are other potential advantages of using these techniques over traditional econometric models. The linear regression model, for example, is constrained by the parametric assumption that the outcome variable must be a linear and additive function of the model's covariates. In general, machine learning techniques do not have such restrictive conditions. Moreover, Mullainthan and Spiess (2017) note that machine learning algorithms are optimized to find non-linear patterns in data, such that they are able to predict the "highly nonlinear and interactive" patterns that are used for facial-recognition problems. Intuitively, we would expect personality traits and soft skills to affect job outcomes in complicated, nonlinear ways, making machine learning an attractive solution to this challenge.

Interpretability is a key concern in choosing a model for our application because the output needs to be comprehensible to a hiring manager. In other words, we need to know why an applicant is matched to a specific job, and what about their psychometric profile makes them a good fit. However, despite the advantages in regards to the predictive accuracy of machine learning models, economists tend to agree that they can be difficult to interpret. Athey and Imbens (2019) observe that, in general, machine learning techniques "come at the expense of one of the concerns that the statistics and econometrics literature have traditionally focused on, namely the ability to do inference." The same features of machine learning models that allow for flexibility also tend to make inference difficult; the fact that these algorithms can fit many different functions means "a greater chance that two functions with very different coefficients can produce similar prediction quality," leading to unstable coefficient estimates (Mullainthan and Spiess 2017). In contrast, estimates in parametric models, such as linear regression and multinomial logistic regression, have estimates that are more stable, and can more intuitively give a sense as to the relative importance of different soft skills in labor outcomes.

While interpretability is key, our model need not have any causal interpretation. In fact, we would actually want to avoid adding covariates that might normally be used as controls in causal studies (such as gender, age, and race) to avoid creating a biased algorithm. Even without using these types of control, Veale and Binns (2017) warn that training decision-making models with data which "reflect existing, unwanted discrimination in society" will cause these models to "encode these same patterns, risking reproduction of past disparities." As this project involves matching applicants to job opportunities, avoiding models that perpetuate implicit and explicit bias will be a crucial consideration.

Data

One of the primary challenges in this project is determining how to measure the "best fit" between a worker and a job. Different workers tend to have different criteria for choosing a job; in general, it is a balancing act between many factors including wages, difficulty of the job, and overall enjoyment or job satisfaction. In addition, employers have their own set of criteria for hiring and keeping employees, especially related to education levels and impressions of workers' soft skills. Seeing as employment is a mutual agreement between employer and worker, using job satisfaction alone as a metric for job fit might be too one-sided for our specific question. For example, it might be the case that a worker has high satisfaction for a job, but is fired after only a few weeks. Better indicators of good worker fit include high wages, long tenure, or high aggregate wages over a worker's tenure at a given job.

Our strategy for addressing this challenge is using longitudinal data on job outcomes and worker characteristics, namely the National Longitudinal Survey of Youth (NLSY97), a study undertaken by the U.S. Bureau of Labor Statistics. This study contains the results of 19 different surveys from 1997 to 2021 from the same sample cohort of 8,984 Americans. The advantage of using longitudinal data is that we can impute the "best" job for an individual worker out of all that worker's job in the 24 years span of this study, given a criteria for job fit. For example, for a worker that had multiple jobs from 1997 to 2021, we might consider their "best" job to be the one in which they had the longest tenure. Ultimately, the criteria we chose is total earnings per job, since this is a function of both wages and tenure. We calculate total earnings for each job (j) for each individual (i) in the sample using the variables shown below in Figure 1. Then, for each individual in the sample we choose the job with the maximum earnings and associate it with one of 17 industry categories, as seen in Appendix 1.

$$\text{Earnings}_{ji} = \text{Tenure}_{ji} * \text{HourlyComp}_{ji} * \text{HoursPerWeek}_{ji}$$

Figure 1: Total Earnings Calculation

In addition to data on labor outcomes, NLSY97 contains survey questions meant to assess various personality traits of the respondents. The Bureau of Labor Statistics does not provide an explanation as to how they chose the 58 survey questions in this section, however most appear to overlap with traits that are commonly used in the psychology literature such as GRIT and “the Big Five.” With the help of documentation from Cahill et al**.**, we grouped the survey questions for NLSY97 into five overarching soft-skill categories: GRIT, agreeableness, conscientiousness, ability to overcome stress, and openness.

To train our machine learning models we use both the raw personality survey questions (hereafter referred to as “disaggregated data”), as well as a separate dataset containing only the scores in each of our five soft-skill categories–our “aggregated” psychometrics– and the indicator dummies for missing data. To create our aggregated psychometrics, we first reverse-code any survey questions that relate negatively to its corresponding soft skill. For example, for a question where a respondent is asked to rank themselves on a scale of one to five for how “[they] feel that undependable describes [them] as a person ("National Longitudinal Surveys.”),” we reverse the scale such that a score of five reflects the presence of a “good” trait or soft skill (in this case conscientiousness), as opposed to a “bad” trait (undependableness). Then for each respondent in the survey, their score in a certain soft skill category is just the sample average of each survey question's response corresponding to the respective category.

NLSY97 contains data for at least one job for nearly every one of the respondents in the 8,984-person sample. Therefore, using one of the job fit criteria mentioned above, we can impute the “best job” for the vast majority of respondents in this sample. However, due to the large amount of non-responses for the personality trait questions, simply dropping rows with “N/A” values would have resulted in having too few data to train any kind of model. Following Taddy (2019) we use zero imputation, in which all missing values are replaced with zero and flagged with a dummy indicator specific to the missing variable.

Empirical Methods

I. Factor Analysis

The goal of this project is to develop and refine a methodology for evaluating survey data of soft skills and personality traits to best match youth to jobs in South Africa and Ethiopia. Since we aim to create a process that systematically yields accurate and predictive results, we must be careful in demonstrating that the underlying framework of measuring and identifying soft skills and matching individuals to jobs is methodologically valid.

According to a large body of psychology literature, the Big Five Personality Traits (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism) and GRIT can reliably define an individuals' personality and soft skills. Given some preliminary analyses and the limitations of our dataset, our analysis will focus on GRIT, Agreeableness, Conscientiousness, Openness, and Stress. Per standards outlined in World Bank documentation, we sorted the NLSY97 survey into these soft skill groups and then performed some basic factor analyses using Stata.

We first standardized each of the survey questions so that the responses aligned in terms of their interpretation. More specifically, questions were realigned so that all survey questions reflected a low score for a low level of the underlying trait. Next we calculated Cronbach's alpha, a measure of how closely related a group of items are, separately for each of the soft skill groups. Again, per World Bank guidance, a measure of 0.7 or higher indicated that a grouping of items is internally consistent. A visualization of calculating Cronbach's alpha, denoted as Test scale, is shown below in Figure 2. The important features shown in this figure are that each survey question is an item, item-rest correlation, each item's alpha, and test scale. Through existing documentation and consensus, items with correlation differing drastically from the others, should be dropped. As shown in Figure 2, the second item that is *Deals w/ setback* meets the criteria of being dropped from the grouping and was dropped before proceeding in the steps of analysis.

GRIT Cronbach's Alpha

Item	Obs	Sign	Item-Test Correlation	Item-Test Correlation	Avg inter-item Covariance	Alpha
New Ideas	3152	+	0.5746	0.3883	0.2244839	0.6751
Deals w/ setback	3152	+	0.333	0.0758	0.283998	**0.7534**
Short Term Obj.	3152	+	0.6502	0.4766	0.2058528	0.6539
Hardworker	3152	+	0.4755	0.3556	0.2560598	0.6857
Changes Goals	3152	+	0.6541	0.4941	0.2076144	0.6506
Stays Focused	3152	+	0.7373	0.5939	0.1854186	0.6237
Finishes Projects	3152	+	0.6521	0.5178	0.2152233	0.6501
Diligence	3152	+	0.5337	0.3604	0.2359154	0.6809
Average					0.2268208	0.7028

Figure 2: Cronbach's Alpha sample calculation for GRIT soft skill grouping

Following the grouping validation, the number of common factors was determined to be able to use the information of the variance of observations on covariates. Horn's Parallel Analysis, as shown in Figure 3 below, yields only one eigenvalue greater than 1. Referring to discussion with our project lead and existing Stata documentation, eigenvalues greater than 1 correspond to significant factors within a grouping.

Factor Analysis/ Correlation

Method: Principal Factors		Obs:	3152
Rotation: (unrotated)		Retained Factors:	1
		Params:	7

Factor	Eigen Value	Difference	Proportion	Cumulative
Factor 1	2.16621	1.174868	1.0944	1.0944
Factor 2	0.41753	0.42032	0.2109	1.3053
Factor 3	-0.00279	0.10939	-0.0014	1.3039
Factor 4	-0.11218	0.00196	-0.0567	1.2473
Factor 5	-0.11414	0.04386	-0.0577	1.1896
Factor 6	-0.158	0.05927	-0.0798	1.1098
Factor 7	-0.21726	-	-0.1098	1.0000

LR test: independent vs. saturated: chi2(21) = **4627.83** Prob>chi2 = **0.0000**

Figure 3: Horn's Parallel Analysis for factor analysis. *Key points are the eigenvalue column: eigenvalues greater than one indicate a significant factor.*

To further understand the information that each item provides within a grouping across possible scores, a graded response model for the item responses of the ordered categorical data is used as shown in Figure 4 below. The items that provide the best information on said soft skill group are not flat in slope and are typically skewed to one side.

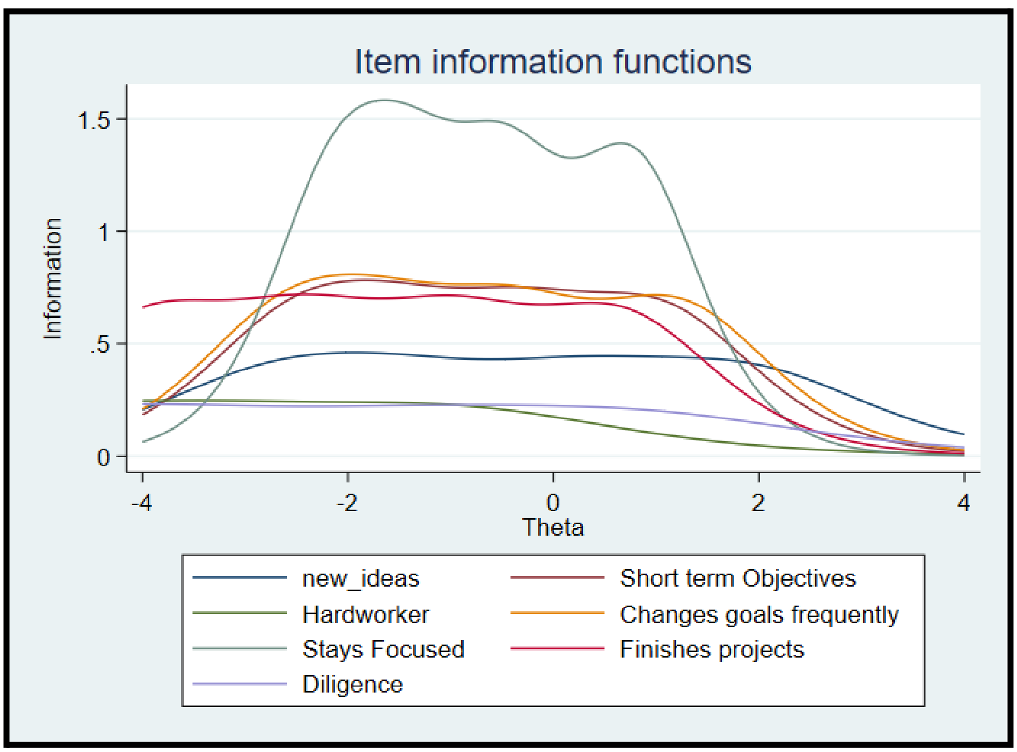


Figure 4: Item information from graded response model.

This is a visualization of how much differences of scores within a survey question inform on the grouping factor. Furthermore, a row mean can be calculated across the variables within said personality grouping and graphed to show the density of total item scores as shown in Figure 5 below. Note, row means are able to be calculated in an accurate manner as missing data resulted in an individuals

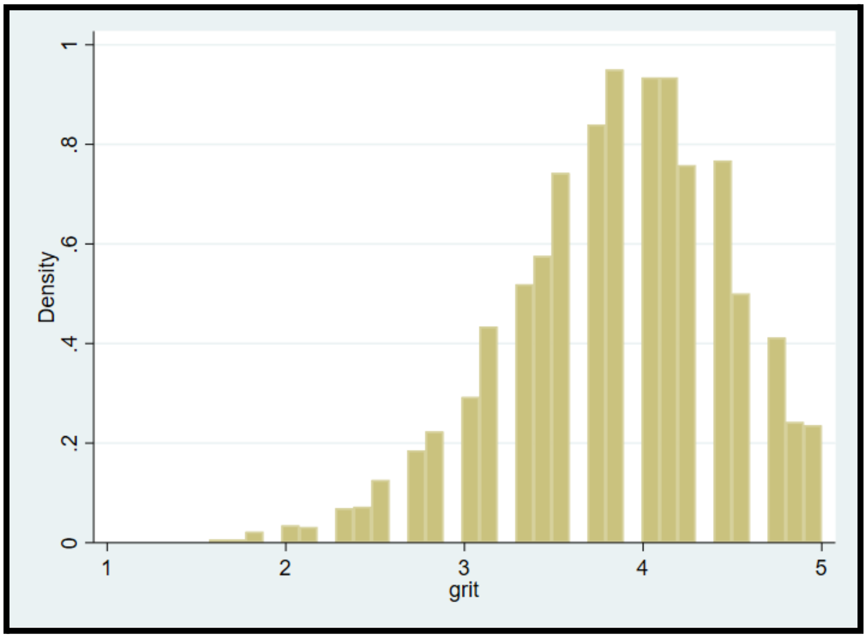


Figure 5: Distribution of total survey response scores for GRIT grouping.

The methodology and framework behind this analysis follows the commonly held consensus of those producing work with soft skills data and the existing documentation within the World Bank for valid results. The issue arises in that the implementation is case by case and requires hands-on work, in other words this approach lacks the flexibility to deal with changes in the data.

II. Machine Learning

For this project we aimed to choose models which, in addition to providing accurate matches, provided useful information about the matches to applicants and employers. With a variety of different classification machine learning tools to choose from, we focus on four that we think could be particularly useful for interpretation: decision trees, random forests, neural networks, and multinomial logit with a regularization parameter.

Decision trees provide generally strong prediction results and allow users to take advantage of several different regularization techniques. Specifically, we utilize cost-complexity pruning (CCP) and experiment with 271 different CCP alphas (the regularization hyperparameter in this case). Another benefit of decision trees is that they have useful interpretations via their use of the Gini Criteria, in which one can see which variables provide the most useful "split" between different units into two homogenous groups.

Random Forest is a useful extrapolation of the Decision Tree, in which many trees are averaged together. With this extra flexibility, Random Forest tends to be more accurate than Decision Tree and, as an extra benefit, comes with built-in features in scikit-learn to show the importance of the covariates (again, based on how well units can be split into homogeneous groups based on a certain feature).

Finally, given the popularity of neural networks in applications ranging from image recognition to causal models, we fit the data with a neural network classifier using the TensorFlow library. Neural networks are useful in a variety of different contexts because of their ability to capture highly non-linear patterns in data through a process of passinging information through layers of nodes or "neurons." Flexible models, though highly useful in prediction problems, are prone to overfitting–to avoid this we utilize dropout layers between each node layer in our model. Dropout layers prevent overfitting by randomly setting node input values to zero throughout the training process.

Appendix 2 demonstrates the importance of using dropout layers and other techniques to reduce overfitting. At each step or "epoch" in the training process, we test the model's accuracy using the training dataset as well as a validation data set. With a highly overfit model, one would expect a large divergence between the training and validation accuracy. The difference between our training and validation accuracy tends to be around two percentage points, however there are large shocks in accuracy which are likely due to the probabilistic nature of implementing the dropout layers.

The multinomial logit in this application lends itself to a more traditional econometric approach. Compared to most machine learning techniques, multinomial logit has useful, well-documented statistical properties which allow for a greater focus on interpretable results–especially with regard to marginal effects. Specifically, in this application, it would be useful to see how marginal effects change for different units in the sample. After all, although there are specific soft skills that employers value, we would not necessarily expect employers to value them equally for all candidates. Still, because we are using a regularization parameter and we do not control for nearly enough covariates to show causal effects, the parameter estimates in our model are biased and we must be careful in interpreting them.

Results

The prediction accuracy of the various models were unsurprisingly low, given the limitation on covariates that could be reasonably added to improve the results. Still, with the worst-performing model predicting the best industry fit for a worker 12.02% percent of the time, each model performed better than if one were to assign jobs randomly (in which case a worker would have a one in seventeen chance–about six percent–of being placed into their best industry). The accuracy results of each model are shown below in Figure 6.

Model Accuracy Percentages

Model	Disaggregated	Aggregated
Decision Tree	16.92	16.53
Neural Network	19.2	16.86
Random Forest	18.14	12.02
Multinomial Logit with Regularization	19.53	18.64

Figure 6: Disaggregated and Aggregated Model Accuracy Results

In addition, as one might expect, the models that were trained with the disaggregated data performed better than those trained with aggregated psychometrics. In general, machine learning models tend to be more accurate if they are trained with more data. Recall that our disaggregated dataset contained the raw responses to each individual survey question in NLSY97. Although having a single metric for each soft skill may be convenient as a summary for employers of an applicant's social skills, it does not provide as much information as a variety of scores measuring different aspects of a certain trait. For example, one intuitively knows that the ability to not get distracted from work is different from the quality of being a hard worker– however these traits are both subsumed under GRIT. Still, for building a platform to match applicants to jobs, the World Bank need not constrain itself to one type of data. Rather, it could make sense to use aggregated data to score applicants on soft skills and use disaggregated data to decide on the best match.

Appendix 3 provides a more in-depth visualization of how our multinomial logit model performs in terms of out-of-sample accuracy for the disaggregated data. The confusion matrix, with rows i and columns j, is defined such that $C_{i,j}$ is equal to the number of observations known to be in group i and predicted to be in group j. The fact that this model, which only conditions on applicants' soft skills, is frequently "confused" about which industry to place applicants has a potentially straightforward explanation: different types of employers tend to value the same soft skills, so applicants with particular soft skill profiles are a good match for several different industries. Kern et al. (2019) find that this holds true even when looking at more specific, job-level data.

Appendix 4 tells a similar story. By default, neural network classifiers from the python library TensorFlow output predictions in terms of probabilistic guesses for each category. For this particular individual, chosen randomly from the disaggregated test data set, we see that there is low confidence that they fit neatly in one particular industry– the two most likely being professional and related occupations, then sales and related occupations. Probabilistic predictions such as these may provide useful information to applicants about different potential job options.

As stated previously, Random Forest provides relatively simple and intuitive feedback into how the algorithm makes classifications through feature importances. Figure 7 shows the top five features for the Random Forest we ran with the disaggregated data. The top most important variables for categorizing individuals to industries in this case are related to rule-following, which we classify as an aspect of conscientiousness. This is not particularly surprising, given that research consistently shows conscientiousness as an important factor in labor outcomes (Sackett, P. R., and P. T. Walmsley 2014) . Interestingly, none of the five soft skills categories we studied showed up in even the top 29 feature importances in the aggregated Random Forest model–rather they were all dummy indicators for missing data. This further highlights the difficulties of training machine learning models with aggregated psychometrics, as opposed to raw survey data.

Personality Scale	
1.) Following Rules	0.031736
2.) Bending Rules	0.030178
3.) Support for Rules and Traditions	0.029413
4.) Reserved, Quiet	0.029237
5.) Critical, Quarrelsome	0.029063

Figure 7: Top 5 Features on Disaggregated Data using Random Forests

Discussion

Our results seem to indicate that we can partially predict job matches using soft skills data, despite the limitations of our methods and data. With factor analysis results consistent with the personality and soft skills literature, we are confident that the skills we identified were properly categorized and measured. Although the prediction accuracy is low, there seems to be evidence that machine learning and specialized econometrics methods can help place disadvantaged workers into jobs based on their soft skills and personality traits. This is very promising for the goals of the World Bank. The results of the disaggregated Random Forest model also provides valuable information. In particular, knowing that "rule following" seems to be a critical skill, policy interventions can better target and place individuals.

Some limitations of our methodology center around the disadvantages of our chosen methodologies and the limitations of the NLSY97 data. While the machine learning prediction accuracy was low, ideally we would be able to increase the accuracy, especially for a human resources end product. Furthermore, there are some external validity concerns with using data from the United States only. Even if our results are relatively accurate there is no guarantee that they apply to the target populations in Ethiopia and South Africa.

Future research in this area should center around expanding the machine learning and econometrics techniques employed as well as explore different datasets. We suggest trying some more complex machine learning models such as causal forests or double machine learning. These techniques could increase prediction accuracy and allow for a causal interpretation of our results. We also recommend finding individual correlations to jobs for better interpretation. For example, analyzing and compiling the data so each individual has a "correlation" for each job (Person A is a 87% match with Sales and Related Occupations, 42% match with Healthcare

Support, etc.). Along with these new models we would recommend strong and diverse robustness checks and sensitivity analyses to further internally validate the results.

Incorporating new datasets from Ethiopia and South Africa for this kind of analysis will likely be a deliberative process that draws from previous work along with some amount of automation for certain tasks. Cahill et al. (2021) demonstrate the challenge of automating the process of grouping personality survey questions into different overarching soft skill categories. They find that unsupervised learning based on clusters of similar words tends to produce groupings which lack subjective interpretability. With the lack of a large labeled dataset mapping survey questions to soft skill categories, supervised machine learning likewise becomes unworkable.

Factor analysis, on the other hand, may lend itself to automation in future projects; researchers would simply need some rule–probably based on some threshold value of Cronbach's alpha– to decide which survey results show internal consistency. Still, the data cleaning which usually proceeds factor analysis will probably require human discretion. After all, reverse-coding survey responses to positively correspond to their soft skill categories is a highly specific task which would require its own sophisticated natural language machine learning model.

Despite these limitations we are confident we have identified a partial link between soft skills and labor market outcomes for the World Bank's Human Resources setting. Our results suggest that the World Bank should focus on rule following to more accurately place and match disadvantaged youth to jobs based on their soft skills. Given the specific limitations of our research, more machine learning and econometric models should be utilized in this area to keep progressing this area of research to increase accuracy and check for robustness of the results.

References

Acosta, Pablo, Noel Muller, and Miguel Alonso Sarzosa. "Beyond qualifications: returns to cognitive and socio-emotional skills in Colombia." *World Bank Policy Research Working Paper* 7430 (2015).

Athey, Susan, and Guido W. Imbens. “Machine Learning Methods That Economists Should Know About.” *Annual Review of Economics*, no. 1, Annual Reviews, Aug. 2019, pp. 685–725. *Crossref*, doi:10.1146/annurev-economics-080217-053433.

Cahill, Ben, Sahil Bobba, Mason Ogden, Jay Ahn, Dennis Sun, and Jonathan Ventura. 2021. *Using Soft Skills to Predict Labor Market Outcomes.*

Cunningham, Wendy, and Paula Villaseñor. "Employer voices, employer demands, and implications for public skills development policy." *World Bank Policy Research Working Paper* 6853 (2014).

Deming, David J. "The growing importance of social skills in the labor market." *The Quarterly Journal of Economics* 132, no. 4 (2017): 1593-1640.

Guerra, Nancy, Kathryn Modecki, and Wendy Cunningham. "Developing social-emotional skills for the labor market: The PRACTICE model." *World Bank Policy Research Working Paper* 7123 (2014).

Jardim, Jacinto, Anabela Pereira, Paula Vagos, Inês Direito, and Sónia Galinha. "The Soft Skills Inventory: developmental procedures and psychometric analysis." *Psychological Reports* 125, no. 1 (2022): 620-648.

Kern, Margaret L., Paul X. McCarthy, Deepanjan Chakrabarty, and Marian-Andrei Rizoiu. "Social media-predicted personality traits and values can help match people to their ideal jobs." *Proceedings of the National Academy of Sciences* 116, no. 52 (2019): 26459-26464.

Mullainathan, Sendhil, and Jann Spiess. “Machine Learning: An Applied Econometric Approach.” *Journal of Economic Perspectives*, no. 2, American Economic Association, May 2017, pp. 87–106. *Crossref*, doi:10.1257/jep.31.2.87.

“National Longitudinal Surveys.” National Longitudinal Surveys | A Program of the U.S. Bureau of Labor Statistics. Accessed April 10, 2022. https://www.nlsinfo.org/.

Sackett, P. R., and P. T. Walmsley. “Which Personality Traits Are Most Important to Employers?” Association for Psychological Science - APS, September 25, 2014. https://www.psychologicalscience.org/news/minds-business/which-personality-traits-are-most-important-to-employers.html.

Taddy, Matt. *Business Data Science: Combining Machine Learning and Economics to Optimize, Automate, and Accelerate Business Decisions*. McGraw-Hill Education, 2019.

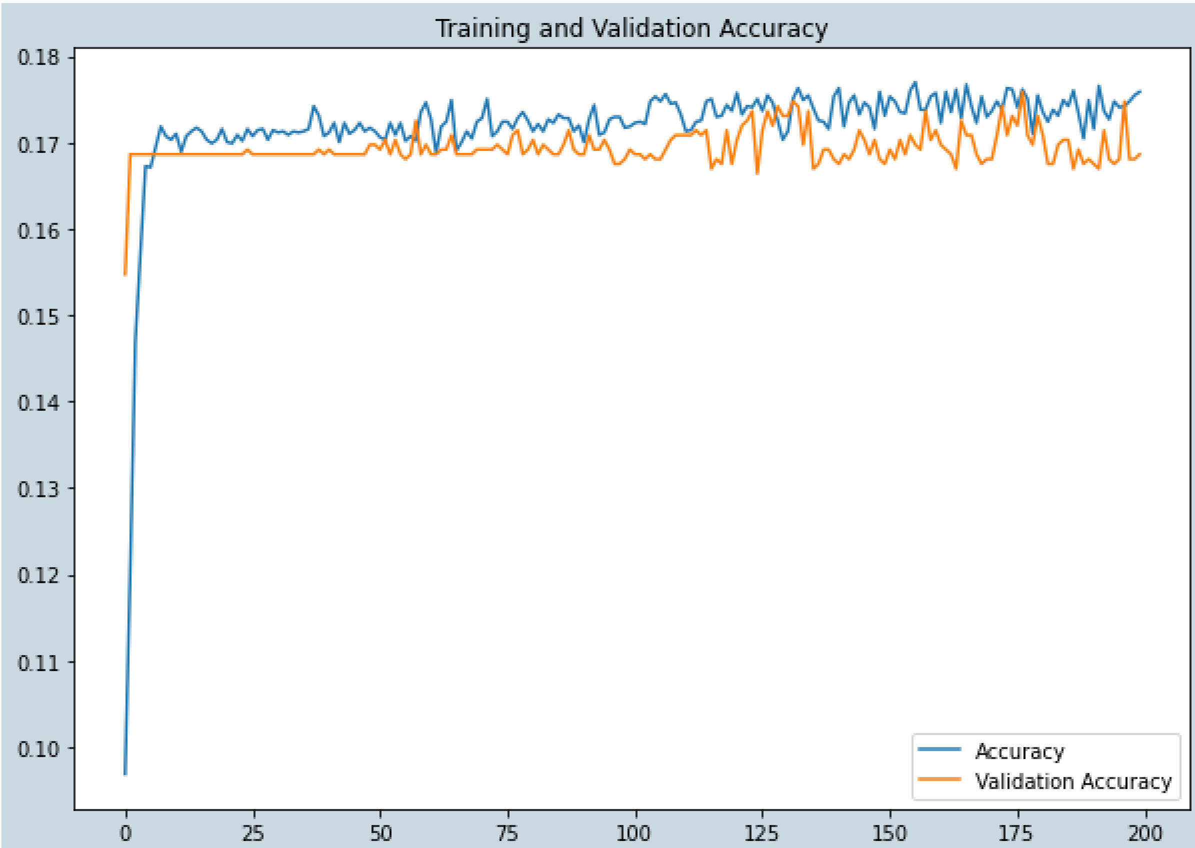
Varian, Hal R. “Big Data: New Tricks for Econometrics.” *Journal of Economic Perspectives*, no. 2, American Economic Association, May 2014, pp. 3–28. *Crossref*, doi:10.1257/jep.28.2.3.

Veale, Michael, and Reuben Binns. “Fairer Machine Learning in the Real World: Mitigating Discrimination without Collecting Sensitive Data.” *Big Data & Society*, no. 2, SAGE Publications, Nov. 2017, p. 205395171774353. *Crossref*, doi:10.1177/2053951717743530.

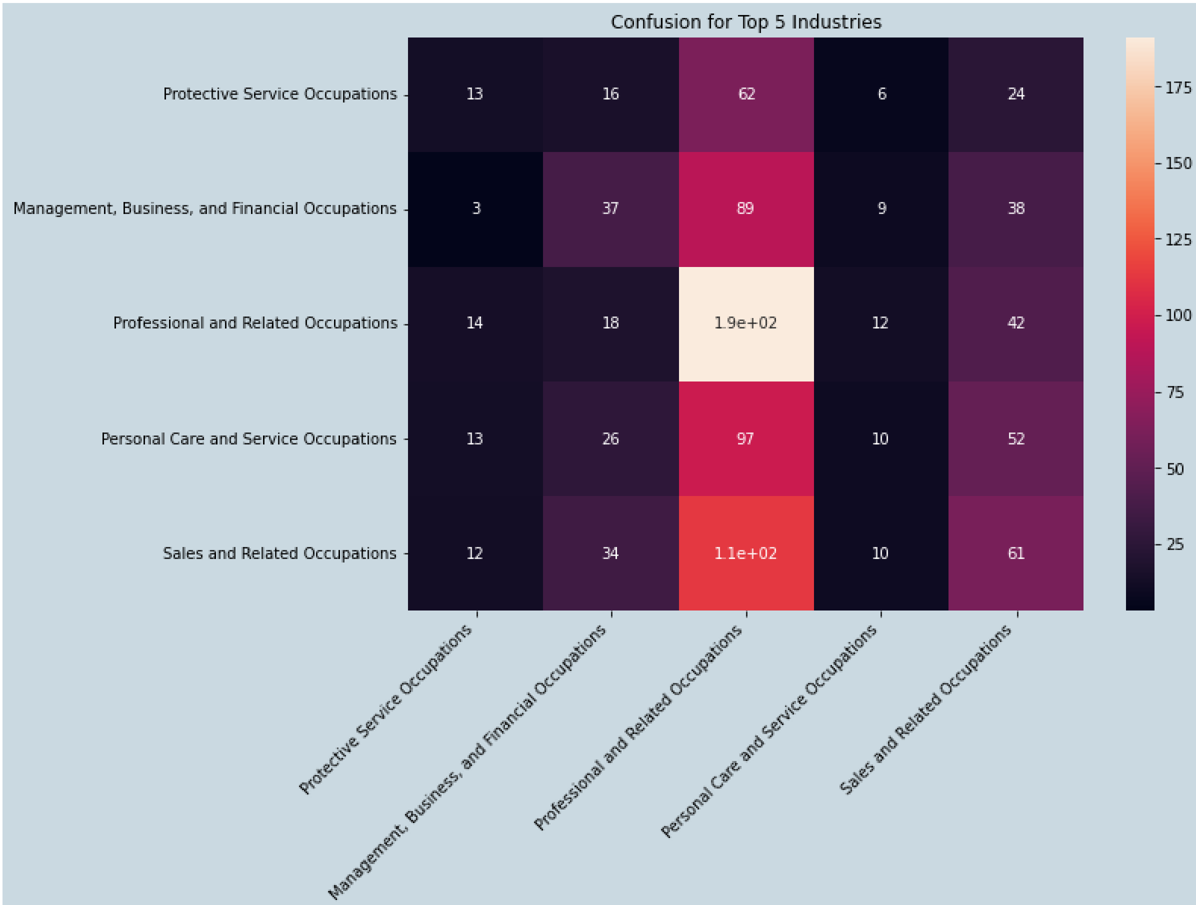
Appendix 1: Industry Categories

1	Management, Business, and Financial Occupations	10	Agriculture Workers including Supervisors
2	Professional and Related Occupations	11	Fishing and Hunting, and Forest and Logging Workers
3	Healthcare Support Occupations	12	Construction and Extraction Occupations
4	Protective Service Occupations	13	Installation, Maintenance, and Repair Occupations
5	Food Preparation and Serving Related Occupations	14	Production Occupations
6	Building and Grounds Cleaning and Maintenance Occupations	15	Transportation and Material Moving Occupations
7	Personal Care and Service Occupations	16	Military
8	Sales and Related Occupations	17	ACS Special Code
9	Office and Administrative Support Occupations		

Appendix 2: Neural Net Classifier Graph



Appendix 3: Confusion Matrix for Top 5 Industries



Appendix 4: Neural Network Probabilistic Prediction for Random Person.