

INTRODUCTION

- Findings from workplace safety research indicate substantial focus on safety climate and less focus on occupational risk factors.
- An integrated approach using machine learning techniques and job analysis information enables novel perspectives on person- and situation-based antecedents to occupation-level safety outcomes.
- RQs: What are the work, worker, and work context features that are most important in predicting safety outcomes? What are their relative contributions?

METHOD

- O*NET job descriptor ratings from job incumbents, analysts to predict nonfatal occupational injury rates published by BLS.
- $N = 376$ data points with a total of 246 predictors from three O*NET domains.
- Three ML models: elastic net (E-Net), random forest (RF), gradient boosting (GB) developed and analyzed using sci-kit learn package in Python v3.7.

RESULTS & DISCUSSION

- 44 unique (18 worker characteristics, 16 work context features, 10 work features) and 12 common predictors across models explained 54.2–58.8% of the variance in accident rates.
- Findings are in line with prior research using traditional statistical methods.
- Suggested exploratory approach helps uncover understudied predictors of accidents such as worker personality, interpersonal dynamics, job working conditions, etc.
- Findings support use of ML techniques to extend findings from traditional statistical methods.
- Identification of critical safety antecedents helps recognize high risk jobs and aids accident prevention through incorporation in safety assessments and training interventions.

Kumar, L.S.¹ & Burns, G.N.²

¹University of South Florida ²Florida Institute of Technology

Advanced machine learning techniques can be used to predict occupational safety outcomes and help uncover understudied predictors.

Machine Learning Model	Elastic Net	Random Forest	Gradient Boosting
# Selected predictors	35	11	17
MSE (Training/Test set)	0.55/0.61	0.14/0.63	0.08/0.68
R ² (Training/Test set)	.679/.588	.916/.570	.955/.542

O*NET Descriptor Category	# Total predictors	# Selected predictors	% Selected predictors
Abilities	52	8	15
Knowledge	33	4	12
Skills	35	3	9
Occupational Interests	6	1	17
Work Values	6	1	17
Work Styles	16	1	6
Generalized Work Activities	41	10	24
Work Context	57	16	28



..... [Link to publication
lkumar1@usf.edu](mailto:lkumar1@usf.edu)

ADDITIONAL INFORMATION

Table 1. Elastic Net Regression Results

Variable	O*NET Descriptor	β	FI (%)
AB_45	Static Strength	0.221	7.256
WC_6	Deal With Physically Aggressive People	0.168	4.395
GWA_29	Performing General Physical Activities	0.121	6.590
AB_43	Speed of Limb Movement	0.117	6.263
IN_4	Investigative	-0.110	2.954
WC_16	Exposed to Minor Burns, Cuts, Bites, or Stings	0.104	4.507
WC_3	Coordinate or Lead Others	-0.090	0.895
SK_32	Technology Design	-0.088	2.467
GWA_35	Resolving Conflicts & Negotiating with Others	0.075	0.915
WC_39	Responsible for Others' Health and Safety	0.073	4.338
WV_6	Working Conditions	-0.068	3.153
KNW_20	Mathematics	-0.068	2.681
AB_11	Far Vision	0.062	2.602
KNW_9	Design	-0.057	1.440
WC_20	Face-to-Face Discussions	-0.055	1.112
WC_32	Level of Competition	-0.050	1.379
WC_30	Indoors, Not Environmentally Controlled	0.046	2.402
WC_49	Spend Time Walking and Running	0.043	5.256
WC_19	Extremely Bright or Inadequate Lighting	0.028	2.845
SK_24	Reading Comprehension	-0.028	3.808
WC_44	Spend Time Kneeling, Crouching, or Crawling	0.025	4.274
GWA_25	Monitoring and Controlling Resources	0.023	0.262
WC_27	In an Enclosed Vehicle or Equipment	0.022	1.956
WC_57	Work With Work Group or Team	-0.021	0.604
GWA_38	Staffing Organizational Units	0.021	0.346
SK_22	Programming	-0.019	3.835
KNW_16	Foreign Language	0.017	0.618
GWA_16	Guiding, Directing, Motivating Subordinates	0.017	0.410
AB_15	Glare Sensitivity	0.011	3.178
WS_3	Analytical Thinking	-0.010	3.017
GWA_31	Processing Information	-0.010	2.234
WC_7	Deal With Unpleasant or Angry People	0.008	1.921
WC_43	Spend Time Keeping or Regaining Balance	0.006	4.460
GWA_1	Analyzing Data or Information	-0.001	2.821
AB_2	Auditory Attention	0.000	2.808

Note. β is standardized regression coefficient, FI is feature importance as computed by relative weight analysis of the sparse predictors.

Table 2. Common Predictors from Three Models

Variable	O*NET Descriptor	E-Net (β)	E-Net FI (%)	RF FI (%)	GB FI (%)
AB_45	Static Strength	0.221	7.26	3.21	3.55
GWA_29	Performing General Physical Activities	0.121	6.59	28.83	24.94
AB_43	Speed of Limb Movement	0.117	6.26	3.41	3.13
WC_6	Deal With Physically Aggressive People	0.168	4.40	1.34	2.51
WC_39	Responsible for Others' Health and Safety	0.073	4.34	-	2.73
SK_22	Programming	-0.019	3.83	-	1.28
WS_3	Analytical Thinking	-0.010	3.02	2.58	3.67
IN_4	Investigative	-0.110	2.95	1.07	1.62
GWA_35	Resolving Conflicts and Negotiating with Others	0.075	0.91	1.06	1.64
AB_47	Trunk Strength	-	-	6.79	9.95
AB_44	Stamina	-	-	2.80	2.81
KNW_7	Computers and Electronics	-	-	2.03	3.30

Note. β is standardized regression coefficient, FI is feature importance as computed by relative weight analysis of the sparse predictors for E-Net and obtained using Python built-in methods for RF and GB.