HUMAN SKILLS PASSPORT: ESTIMATIVA NEURAL DA DEMANDA DE HABILIDADES HUMANAS POR OCUPAÇÃO PARA O RECRUTAMENTO BASEADO EM HABILIDADES

Júlia Alves da Rocha[1]

[1]*Universidade Federal de Alfenas (UNIFAL-MG), Alfenas, Brasil*
(julia.rocha@sou.unifal-mg.edu.br)

Resumo: O *Human Skills Passport* é um sistema de certificação que vincula cada credencial de habilidade humana à evidência comportamental que a originou. Este trabalho avalia seu módulo de match: uma rede neural multitarefa que estima, a partir do texto ocupacional, a demanda de cinco habilidades humanas, treinada com 3,4 mil textos das bases O*NET e ESCO, avaliada em 879 ocupações com rótulos-ouro.

Palavras-chave: redes neurais artificiais; habilidades humanas; recursos humanos; O*NET; contratação baseada em habilidades.

INTRODUÇÃO

Imagine que, durante uma entrevista de emprego, um candidato afirme possuir excelente capacidade de liderança. Como o recrutador pode verificar se essa afirmação é verdadeira? E, mais importante, como pode compará-la de forma objetiva com as declarações feitas por outros candidatos? No processo seletivo tradicional, essa validação é, em grande parte, inviável. Os principais mecanismos utilizados para avaliar competências, como a autodeclaração em currículos, a inferência por proxies, a exemplo da instituição de ensino ou das experiências profissionais, e as entrevistas não estruturadas, apresentam limitações significativas.

Enquanto a autodeclaração carece de mecanismos de verificação, a utilização de proxies introduz vieses e pressupostos indiretos, e as entrevistas não estruturadas apresentam baixa confiabilidade e reduzida consistência entre avaliadores (Dipboye, 1994; Huffcutt e Arthur, 1994; Sackett et al., 2022). Como consequência, competências comportamentais e transferíveis, como liderança, comunicação e colaboração, permanecem de difícil mensuração, comprometendo a qualidade e a equidade das decisões de recrutamento.

Nos últimos anos, observa-se um movimento consistente do mercado de trabalho em direção às habilidades, em detrimento das credenciais formais (Sigelman et al., 2024; Bechtold et al., 2025; OECD, 2025), e, dentro delas, às habilidades humanas de comunicação, colaboração, adaptabilidade, liderança e agilidade de aprendizagem, apontadas como as que mais ganham valor na era da inteligência artificial (McKinsey Global Institute, 2025; World Economic Forum, 2025). Ao mesmo tempo, a literatura de sinalização mostra que empregadores reagem a sinais de habilidade nem sempre confiáveis (Baert et al., 2020), e que frameworks de competências ainda são difíceis de operacionalizar (Ali et al., 2021).

Diante disso, o trabalho foi concebido como um sistema de certificação em que cada credencial de habilidade humana permanece vinculada à evidência comportamental que a originou. Sua arquitetura articula dois módulos. O *badge* é a credencial verificável, emitida a partir de avaliação estruturada e não de autodeclaração. O match estima a demanda de cada habilidade humana por ocupação e informa a aproximação entre candidatos e vagas. Este artigo avalia apenas o segundo módulo. O *badge* é descrito somente na medida necessária para situar o problema, e sua validação não integra o escopo deste trabalho.

O núcleo do *match* é uma rede neural do tipo perceptron multicamadas (MLP) multitarefa: as descrições ocupacionais são convertidas em embeddings de sentenças e a rede aprende, simultaneamente para os cinco construtos, a prever a importância de cada habilidade, utilizando como rótulos os dados da base O*NET (U.S. Department of Labor, 2026). O desempenho é comparado a um baseline de palavras-chave.

O artigo está organizado da seguinte forma: a Introdução contextualiza as limitações dos métodos tradicionais de avaliação de habilidades humanas; a seção Material e Métodos descreve as fontes de dados, a construção do conjunto de dados, a representação textual e a arquitetura da rede neural; a seção Resultados e Discussão apresenta o desempenho do modelo em tarefas de regressão e

classificação, discute os resultados obtidos com o SOM e a Fuzzy ART e expõe as limitações do estudo; por fim, a Conclusão sintetiza os principais achados, as contribuições do trabalho e as perspectivas para pesquisas futuras.

MATERIAL E MÉTODOS

Esta seção apresenta os materiais e métodos empregados neste estudo, incluindo as bases de dados, a preparação do conjunto de treinamento, a representação dos textos ocupacionais, a arquitetura dos modelos neurais e o protocolo de avaliação utilizados.

Fontes de dados e seus papéis no sistema

O treinamento supervisionado utiliza exclusivamente a base O*NET (U.S. Department of Labor, 2026), com 879 ocupações anotadas com a importância de cada um dos cinco construtos: a única fonte disponível com rótulos atribuídos por metodologia ocupacional estabelecida, requisito do aprendizado supervisionado.

As demais fontes compõem a camada de fundamentação de dados do sistema, com função de calibração: o corpus LinkedIn Job Postings 2023-2024 (Kaggle), com cerca de 124 mil anúncios reais, ancora a linguagem das vagas e as estatísticas descritivas de demanda; a taxonomia Lightcast Open Skills normaliza menções de habilidades do texto livre para um vocabulário canônico; a ESCO fornece o mapeamento multilíngue dos construtos, incluindo o português, e, nesta versão, também rótulos prata transferidos pelo crosswalk para uso exclusivo no treinamento; e a API Adzuna mantém o *corpus* de vagas atualizado. Dessas fontes deriva o crosswalk Lightcast-ESCO-O*NET, que constitui o modelo canônico de habilidades da plataforma e origina os rótulos prata da ESCO; por não ter sido validado externamente neste trabalho, eventuais erros de correspondência se propagam para o treinamento sem serem captados pelas métricas, calculadas apenas sobre os rótulos-ouro do O*NET (EUROPEAN COMMISSION, 2022).

O corpus de anúncios reais interage com o modelo neural em dois pontos: o vocabulário das vagas informa o baseline de palavras-chave contra o qual a rede é comparada, e as vagas reais em português e inglês constituem o material da validação qualitativa. Anúncios reais não entram como exemplos de treino por não possuírem rótulos de demanda por construto, decisão metodológica cujo desdobramento natural, a rotulagem humana de uma amostra de vagas para ajuste fino do modelo, é apontada como trabalho futuro.

Construção do conjunto de dados

O conjunto de dados combina uma fonte de rótulos-ouro e uma de rótulos-prata, totalizando 3.455 textos ocupacionais. A fonte-ouro é a base O*NET (versão 29.1; U.S. Department of Labor, 2026): para cada uma de 879 ocupações, o texto de entrada reúne o título, a descrição oficial e até 16 enunciados de tarefas, com prioridade para as tarefas classificadas como essenciais (*Core*), e os rótulos derivam das avaliações de importância (escala IM) dos elementos das tabelas Skills e Work Styles; cada construto é definido por um conjunto fixo de elementos O*NET, mapeados por um crosswalk documentado no código, e seu rótulo é a média das importâncias desses elementos, normalizadas do intervalo 1-5 para 0-1. A fonte prata é a taxonomia ESCO (European Commission, 2024): 2.576 descrições ocupacionais, em inglês e em português, cujos rótulos são transferidos pelo mesmo *crosswalk*; por serem rótulos derivados, e não medidos, esses exemplos participam exclusivamente do treinamento. Cada ocupação é associada ao seu grande grupo SOC (dois primeiros dígitos do código), chave de agrupamento da validação, e o arquivo gerado é identificado pelo seu hash SHA-256, registrado no cartão do modelo.

Representação do texto de entrada

As descrições ocupacionais são convertidas em vetores por um codificador de sentenças multilíngue congelado (*paraphrase-multilingual*-MiniLM-L12-v2; 384 dimensões), o mesmo adotado nos demais módulos do sistema (Reimers e Gurevych, 2019). Como a janela do codificador é de 128 tokens, o embedding é calculado em blocos: um bloco para título e descrição e blocos subsequentes para pares de tarefas, combinados pela média com normalização L2.

O conjunto de treino é expandido por reamostragem de tarefas: cada ocupação gera três variantes aumentadas, além da canônica, cada uma mantendo entre 50% e 85% dos blocos de tarefas, e toda variante herda o fold da ocupação de origem, impedindo vazamento entre treino e teste.

Rede principal: Perceptron Multicamadas Multitarefa

O perceptron multicamadas (*Multilayer Perceptron,* MLP) é uma rede neural direta em que camadas sucessivas de neurônios aplicam uma transformação linear seguida de uma ativação não linear, com pesos ajustados pela retropropagação do erro (Rumelhart et al., 1986); com uma camada oculta suficientemente larga, a arquitetura é um aproximador universal de funções contínuas (Cybenko, 1989). Por essas propriedades, é a escolha de referência para regressão e classificação sobre representações vetoriais de

dimensão fixa, como os *embeddings* descritos anteriormente.

No sistema, o MLP cumpre a função de cabeça de demanda do módulo de *match*: recebe o embedding do texto ocupacional e devolve, em uma única passada, a intensidade de demanda dos cinco construtos. É a única tarefa do módulo com rótulos numéricos disponíveis para o aprendizado supervisionado e substitui o analisador léxico anterior nos construtos expressos de forma indireta no vocabulário das tarefas.

A instância adotada tem uma camada oculta de 128 neurônios com ativação ReLU (Nair e Hinton, 2010; He et al., 2015) e cinco saídas sigmoides, uma por construto, sobre a entrada de 384 dimensões. A largura da camada oculta não resultou de busca sistemática de hiperparâmetros, mas de decisão de projeto orientada por parcimônia e regularização: dado o volume limitado de ocupações rotuladas disponíveis, uma camada estreita restringe a capacidade do modelo e mantém a razão entre número de parâmetros e volume de dados em faixa favorável à generalização. A otimização sistemática de largura, profundidade e taxa de *dropout*, por validação cruzada estratificada, fica reservada a etapas posteriores, condicionada à obtenção de volume de dados que a justifique.

O tronco é compartilhado entre os construtos em formulação multitarefa, na qual os construtos, correlacionados entre si, atuam como regularização mútua, e o total de parâmetros permanece na ordem de dezenas de milhares. O treinamento minimiza a entropia cruzada binária (BCE) sobre os rótulos contínuos em 0-1 com o otimizador AdamW (Loshchilov e Hutter, 2019) e parada antecipada monitorada em uma fatia de validação interna, também agrupada por setor.

O resultado final é um ensemble de cinco sementes cuja saída é a média dos logits, reduzindo a variância de inicialização; os hiperparâmetros completos e o hash do conjunto de dados ficam registrados no cartão do modelo. O modelo é exportado em formato ONNX com contrato fixo de entrada (embedding, 1×384) e saída (logits, 1×5), o que permitiu substituir o analisador anterior sem alteração da interface de serviço.

Baseline de comparação

O baseline é o analisador de palavras-chave determinístico originalmente embarcado no módulo de *match*: um crosswalk léxico bilíngue (português e inglês) que associa termos e expressões de vagas a cada construto e pontua a demanda pela frequência dessas ocorrências no texto. O baseline recebe exatamente os mesmos textos ocupacionais avaliados pela rede, de modo que a comparação isola o método de estimativa.

Protocolo de avaliação

A avaliação segue a validação cruzada de 5 *folds* agrupada por grande grupo SOC (GroupKFold): em cada rodada, o modelo é treinado sem acesso a nenhum setor presente no fold de teste, e todas as métricas são calculadas apenas sobre previsões *out-of-fold*, na variante canônica do texto (sem aumento). Na formulação de regressão, reportam-se a correlação de Spearman entre demanda prevista e importância observada, o erro absoluto médio (MAE) e o R^2 por construto; na formulação de classificação, define-se demanda alta como valor acima da mediana do construto, balanceando as classes e fixa o acaso em 50%, e reportam-se acurácia, precisão, *recall*, F1 e ROC-AUC. As mesmas métricas são calculadas para o baseline.

Reprodutibilidade: sementes fixadas para Python, NumPy e PyTorch, algoritmos determinísticos habilitados, e cada execução registra a configuração, o hash do conjunto de dados e as métricas por época; os resultados reportados correspondem à média do *ensemble* de cinco sementes.

Mapa auto-organizável (SOM)

O mapa auto-organizável (SOM) é uma rede neural não supervisionada de aprendizado competitivo que projeta dados de alta dimensão sobre uma grade bidimensional de unidades, preservando a topologia do espaço original: entradas semelhantes ativam unidades vizinhas no mapa (Kohonen, 1990). O treinamento aproxima iterativamente os pesos da unidade vencedora e de sua vizinhança de cada entrada apresentada, sem utilizar rótulo algum.

No sistema, o SOM cumpre função de validação exploratória da premissa de que os embeddings do texto ocupacional carregam estrutura suficiente para aprender demanda de habilidades. Se ocupações do mesmo setor se agrupam no mapa sem qualquer supervisão, essa estrutura existe de fato na representação, e não é um artefato dos rótulos. O mapa treinado também alimenta a superfície de exploração visual do espaço ocupacional na plataforma.

Foram comparadas grades de 10×10, 15×15 e 20×20 unidades por dois critérios complementares: o erro de quantização (QE), distância média de cada entrada à sua unidade vencedora, e o erro topográfico (TE), fração de entradas cujas duas unidades mais próximas não são vizinhas no mapa. A qualidade do mapa final é resumida pela U-Matrix (distâncias entre unidades vizinhas, que revela fronteiras de densidade), pelo mapa de acertos e pela pureza setorial, a fração de ocupações atribuídas a unidades

cujo rótulo majoritário coincide com o seu grande grupo SOC.

Monitor de integridade das entradas (*Fuzzy* ART)

A *Fuzzy* ART é uma rede da teoria da ressonância adaptativa para categorização incremental de padrões analógicos: cada entrada é comparada aos protótipos das categorias existentes e, se a semelhança supera o parâmetro de vigilância, a categoria correspondente ressoa e seu protótipo é atualizado; caso contrário, uma nova categoria é criada. A codificação por complemento normaliza as entradas e evita a proliferação descontrolada de categorias (Carpenter et al., 1991). A arquitetura resolve o dilema estabilidade-plasticidade: aprende padrões novos sem apagar os já consolidados, o que a torna adequada à operação contínua.

No sistema, a *Fuzzy* ART atua como monitor de integridade das entradas do módulo de *match*: treinada nos embeddings das ocupações, ela sinaliza como anômala, em produção, a entrada que não ressoa com nenhuma categoria existente (uma vaga malformada, um texto de outro domínio, uma descrição corrompida), encaminhando-a para revisão humana antes que alcance o modelo de demanda. A escolha decorre exatamente das propriedades anteriores: detecção de novidade é o comportamento nativo da rede, e o custo incremental por entrada é baixo o bastante para monitoramento contínuo.

O parâmetro de vigilância governa o compromisso entre sensibilidade e falso alarme, e por isso foi varrido de 0,75 a 0,90 em um conjunto balanceado com 176 exemplos normais e 176 anomalias injetadas, medindo recall de detecção, taxa de falsos positivos, precisão, F1 e o número de categorias formadas.

RESULTADOS E DISCUSSÃO

Esta seção apresenta, primeiro, o desempenho do módulo neural de demanda nas formulações de regressão e de classificação, sempre em comparação com o baseline de palavras-chave e sob previsões out-of-fold; em seguida, a generalização setorial; e, por fim, duas análises complementares com redes não supervisionadas.

Regressão: ordenação da demanda por construto

A Tabela 1 e a Figura 1 apresentam o resultado central. O modelo neural alcança correlação de Spearman média de 0,520 entre demanda prevista e importância observada, contra 0,233 do baseline de palavras-chave, ganho médio de +0,286, com vantagem em todos os cinco construtos.

A hierarquia de desempenho é estável: Comunicação (ρ = 0,686), Colaboração (0,617) e Agilidade de aprendizagem (0,581) no topo; Liderança (0,453) intermediária; Adaptabilidade (0,261) no piso. O contraste é mais revelador onde o construto se expressa de forma indireta no texto: em Liderança, a contagem de palavras-chave produz correlação negativa (-0,037), enquanto a rede recupera o sinal presente no vocabulário de coordenação e supervisão das tarefas. Os erros absolutos contam a mesma história por outro ângulo: o MAE do modelo neural fica entre 0,054 e 0,074, enquanto o do baseline varia de 0,369 a 0,674, mostrando que o método léxico erra na ordenação das ocupações e no próprio nível da demanda.

Tabela 1. Métricas de regressão por construto (*out-of-fold*, rótulos-ouro O*NET): modelo neural vs. baseline de palavras-chave.

Construto	ρ neur.	ρ base.	R^2 neur.	MAE neur.	MAE base.
COM	0,686	0,513	0,474	0,074	0,369
CLB	0,617	0,436	0,355	0,054	0,478
ADA	0,261	0,109	0,007	0,074	0,674
LID	0,453	-0,037	0,205	0,068	0,578
AGI	0,581	0,146	0,315	0,070	0,548

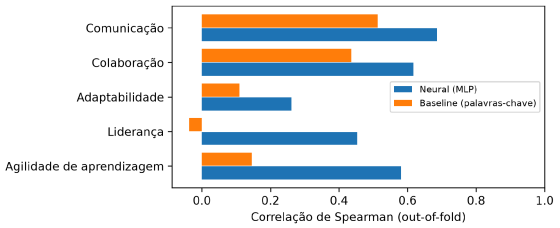

Figura 1. Correlação de Spearman por construto: modelo neural vs. baseline de palavras-chave (*out-of-fold*). Fonte: elaborado pelos autores.

Para evitar que o desempenho reportado reflita uma inicialização favorável isolada, o protocolo de avaliação prevê o treinamento do modelo sob cinco sementes aleatórias distintas (42–46), mantendo fixos a arquitetura, os hiperparâmetros e a partição treino/validação. Os resultados são reportados como média e desvio-padrão da métrica de concordância quadrática ponderada (QWK) por construto, acompanhados dos valores extremos observados. Na validação do pipeline sobre dados sintéticos, o QWK global apresentou média de 0,442 (DP = 0,053; amplitude de 0,370 a 0,500). No nível dos construtos, contudo, a dispersão mostrou-se substancialmente maior, com desvios-padrão entre 0,061 (LEA-1) e 0,217 (COL-2); neste último, o QWK variou de 0,037 a 0,578 entre sementes, sob configuração idêntica. Esse contraste evidencia que a agregação entre dimensões atenua a variabilidade, que a estabilidade aparente da métrica global não se transfere ao nível de construto, no qual cada cabeça de saída é estimada a partir de um subconjunto reduzido de exemplos. Conclui-se que, no regime de dados limitados, o reporte de execuções isoladas por

construto não é metodologicamente sustentável, reforçando a necessidade de resultados agregados sob múltiplas sementes.

Em relação à versão treinada apenas com a O*NET (Spearman médio de 0,559 sob o mesmo protocolo), a incorporação dos 2.576 exemplos prata da ESCO manteve o desempenho médio sobre os rótulos-ouro em patamar equivalente, com dois efeitos qualitativos: reduziu a falha no setor de pior generalização (Serviço social, de -0,087 para -0,026) e passou a expor o modelo, já no treinamento, a descrições ocupacionais em português.

Classificação: identificando ocupações de alta demanda

Na formulação de classificação, os F1 médios dos dois métodos são quase idênticos (0,683 do modelo neural contra 0,681 do baseline), e essa igualdade é instrutiva justamente por ser enganosa. A Tabela 3 mostra como o baseline obtém seu F1: em quatro dos cinco construtos, recall de 1,0 com precisão em torno de 0,50, isto é, o analisador de palavras-chave classifica praticamente todas as ocupações como de alta demanda, um comportamento degenerado que o F1 mascara em classes balanceadas.

As métricas que penalizam esse colapso o expõem: a acurácia do baseline fica em torno do acaso (0,501 a 0,507 nesses construtos, contra 0,681 em média do modelo neural, Tabela 2) e o ROC-AUC médio cai para 0,595, contra 0,756 do modelo neural. Como o módulo de *match* ordena vagas por compatibilidade em vez de apenas rotulá-las, o AUC é a métrica mais próxima do uso real, e é nela que a distância entre os métodos aparece por inteiro (Figura 3: de 0,63 em Adaptabilidade a 0,84 em Comunicação).

Tabela 2. Métricas de classificação do modelo neural (alta demanda = acima da mediana).

Constr.	Acur.	Precisão	Recall	F1	ROC-AUC
COM	0,766	0,775	0,761	0,768	0,839
CLB	0,713	0,718	0,712	0,715	0,795
ADA	0,568	0,575	0,567	0,571	0,630
LID	0,672	0,673	0,673	0,673	0,731
AGI	0,686	0,689	0,686	0,687	0,783

Tabela 3. Métricas de classificação do baseline de palavras-chave (mesmo corte).

Constr.	Acur.	Precisão	Recall	F1	ROC-AUC
COM	0,684	0,652	0,815	0,724	0,748
CLB	0,505	0,505	1,000	0,671	0,672
ADA	0,507	0,507	1,000	0,673	0,545
LID	0,501	0,501	1,000	0,667	0,483
AGI	0,503	0,503	1,000	0,669	0,526

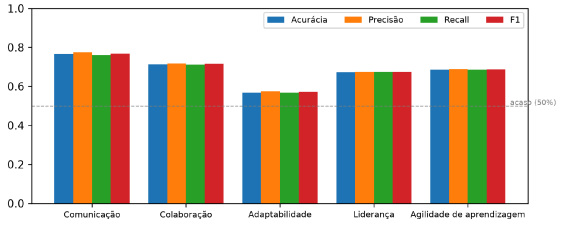

Figura 2. Métricas de classificação de alta demanda (acima da mediana), modelo neural. Fonte: elaborado pelos autores.

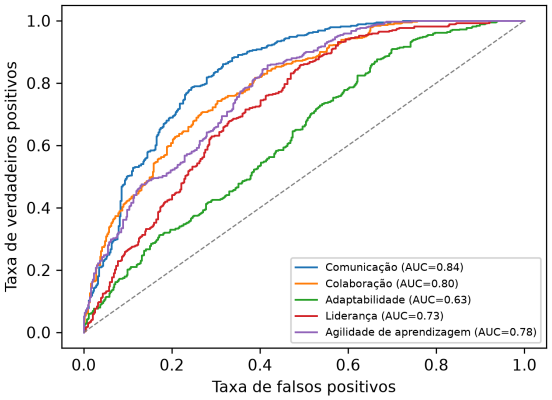

Figura 3. Curvas ROC por construto (classificação de alta demanda, *out-of-fold*). Fonte: elaborado pelos autores.

As matrizes de confusão (Figura 4) complementam o quadro: os erros do modelo neural são aproximadamente simétricos entre falsos positivos e falsos negativos em todos os construtos, sem viés de superestimar ou subestimar a demanda. Em Adaptabilidade concentra-se a maior massa fora da diagonal (380 dos 879 casos), coerente com o diagnóstico da próxima subseção.

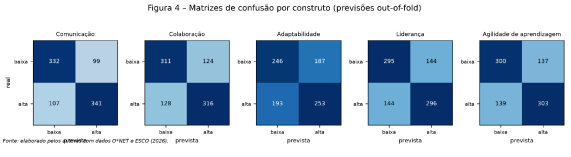

Figura 4. Matrizes de confusão por construto (previsões *out-of-fold*). Fonte: elaborado pelos autores.

Diagnóstico: compressão das previsões e limite dos rótulos

Os diagramas de dispersão entre demanda prevista e observada (Figura 5) revelam compressão das previsões em torno da média: a rede ordena as ocupações corretamente, mas em amplitude menor que a dos rótulos. Isso explica a convivência de erros absolutos baixos (MAE entre 0,054 e 0,074) com R^2 heterogêneo (0,007 a 0,474) e reforça que a propriedade preservada pelo modelo é a ordenação,

não a calibração. Cabe explicitar o limite inferior dessa faixa: em Adaptabilidade, um R^2 de 0,007 significa que o modelo praticamente não explica variância além da média, e suas saídas não devem ser interpretadas como estimativas pontuais de demanda nesse construto. A limitação, porém, antecede o modelo: os rótulos O*NET de Adaptabilidade têm variância muito baixa entre ocupações, de modo que quase todas demandam adaptabilidade em grau alto e semelhante, como a Figura 6 tornará visível, restando pouco sinal disponível para aprender.

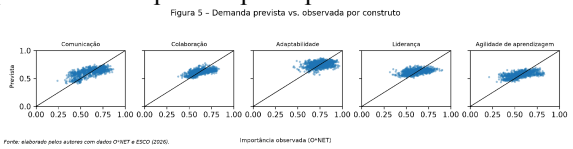

Figura 5. Demanda prevista vs. observada por construto. Fonte: elaborado pelos autores.

Padrões setoriais e generalização

O perfil médio de demanda por setor (Figura 6) funciona como teste de validade de face: serviço social, educação, gestão e saúde concentram as principais demandas de habilidades humanas, enquanto produção, agropecuária e construção apresentam as menores, hierarquia compatível com a literatura recente sobre o valor dessas habilidades (McKinsey Global Institute, 2025; World Economic Forum, 2025), e a demanda de Adaptabilidade é alta em praticamente todos os setores, materializando a baixa variância discutida acima.

Já a Figura 7 decompõe a generalização: como a validação cruzada é agrupada por grande grupo SOC, cada *fold* avalia o modelo em setores inteiramente ausentes do treino. A generalização é mais forte em Transporte/Logística (ρ médio = 0,626; n = 47), Agropecuária (0,584) e Vendas (0,521), e falha em Serviço social (-0,026; n = 14) e Arquitetura/Engenharia (0,083). O primeiro, um setor pequeno e com perfil de demanda atipicamente alto em todos os construtos, difícil de extrapolar; o segundo, um setor cujo texto ocupacional é dominado por vocabulário técnico pouco informativo sobre habilidades humanas.

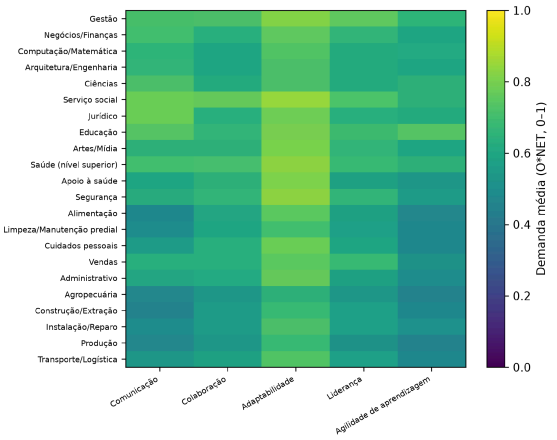

Figura 6. Perfil médio de demanda de habilidades humanas por setor ocupacional. Fonte: elaborado pelos autores.

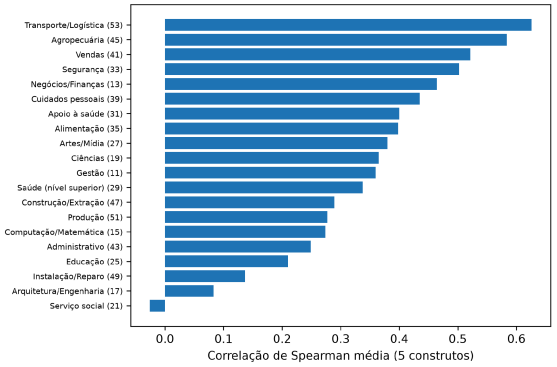

Figura 7. Generalização por setor: correlação de Spearman média em grupos SOC não vistos no treino. Fonte: elaborado pelos autores.

Organização topológica do espaço ocupacional (SOM)

A varredura de grades do mapa auto-organizável (Tabela 4) mostra que a grade de 20×20 unidades domina as demais nos dois critérios: o erro de quantização cai de 0,726 (10×10) para 0,608 e o erro topográfico, de 0,154 para 0,042, com 2,2 amostras por unidade. No mapa final (Figura 8), a U-Matrix exibe fronteiras nítidas de distância entre regiões, e a pureza setorial de 0,81 indica que 81% das ocupações são atribuídas a unidades cujo rótulo majoritário coincide com o seu próprio setor. O resultado corrobora a premissa central de que os *embeddings* do texto ocupacional carregam estrutura setorial suficiente para uma rede aprender demanda de habilidades a partir deles.

Tabela 4. SOM: erros de quantização (QE) e topográfico (TE) por tamanho de grade.

Grade	Unidades	QE	TE	Amostras/unid.
10×10	100	0,7264	0,1536	8,8

15×15	225	0,6595	0,1024	3,9
20×20	400	0,6077	0,0421	2,2

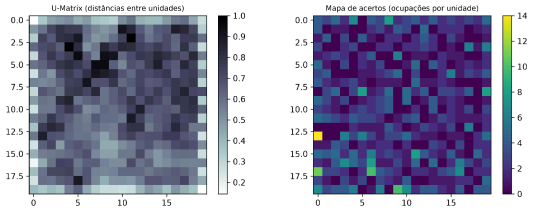

Figura 8. SOM 20×20 sobre embeddings ocupacionais: U-Matrix e mapa de acertos (pureza setorial = 0,81). Fonte: elaborado pelos autores.

Monitor de integridade das entradas (Fuzzy ART)

A varredura do parâmetro de vigilância da rede *Fuzzy* ART (Tabela 5, Figura 9) revela um ponto de operação bem definido. Com vigilância 0,75, metade das anomalias injetadas passa sem detecção (*recall* 0,50); com 0,85 ou mais, o número de categorias explode e mais da metade das entradas normais é sinalizada indevidamente (taxa de falsos positivos de 0,545 a 0,750).

Em vigilância 0,80, a rede detecta 99,4% das anomalias com 14,8% de falsos positivos e F1 de 0,928, desempenho adequado ao papel de monitor de integridade: sinalizar, em produção, entradas fora da distribuição de treino (vagas malformadas, textos de outro domínio) para revisão humana antes que alcancem o módulo de *match*.

Tabela 5. *Fuzzy* ART: varredura de vigilância, detecção de anomalias injetadas (n normal = 176; n anômalo = 176).

Vigilância	Categ.	Recall	FPR	Precisão	F1
0,75	268	0,500	0,000	1,000	0,667
0,80	347	0,994	0,148	0,871	0,928
0,85	373	1,000	0,545	0,647	0,786
0,90	630	1,000	0,750	0,571	0,727

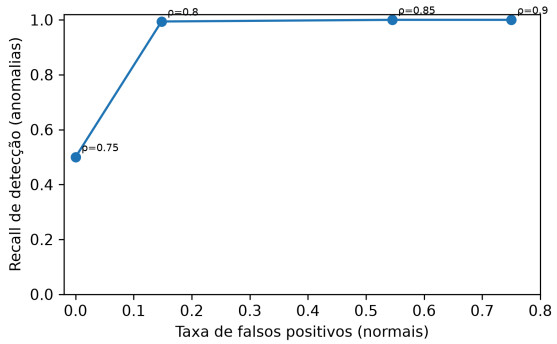

Figura 9. *Fuzzy* ART: compromisso entre detecção e falsos positivos por vigilância. Fonte: elaborado pelos autores.

Limitações

Três limitações delimitam o alcance destes resultados. Primeira: os rótulos-ouro provêm de perfis ocupacionais padronizados, e não de anúncios reais de vagas; a validação com vagas reais em português e inglês permanece qualitativa nesta versão. Segunda: os rótulos ESCO são transferidos por crosswalk, razão pela qual foram confinados ao treinamento. Terceira: em Adaptabilidade, a baixa variância dos rótulos limita estruturalmente o sinal de ordenação disponível, e nenhum aumento de dados de entrada resolve essa restrição; apenas rótulos de outra natureza (por exemplo, anotação humana de vagas reais) poderiam fazê-lo. Coerentemente com o desenho do sistema, o módulo informa a aproximação entre candidatos e vagas, mas não decide autonomamente sobre pessoas; decisões permanecem ancoradas em rubricas comportamentais e revisão humana.

CONCLUSÃO

Este trabalho avaliou o módulo neural de estimativa de demanda de habilidades humanas do *Human Skills Passport*, proposto como resposta à lacuna de verificação que caracteriza essas habilidades nos processos seletivos. Um perceptron multicamadas multitarefa, treinado sobre *embeddings* de texto ocupacional com 3.455 exemplos das bases O*NET e ESCO e avaliado por validação cruzada agrupada por setor em 879 ocupações com rótulos-ouro, superou o analisador de palavras-chave em todos os cinco construtos: a correlação de Spearman média subiu de 0,233 para 0,520, o erro absoluto médio caiu quase uma ordem de grandeza e o ROC-AUC médio passou de 0,595 para 0,756. A comparação também expôs o modo de falha do método léxico, que classifica quase todas as ocupações como de alta demanda, o que o F1 mascara e as demais métricas revelam; o ganho da rede se concentra justamente nos construtos expressos de forma indireta no texto, como Liderança.

O mapa auto-organizável confirmou, sem uso de rótulos, que os *embeddings* ocupacionais preservam estrutura setorial (pureza de 0,81), sustentando a premissa da abordagem; e a rede *Fuzzy* ART, operando em vigilância de 0,80, detectou 99,4% das anomalias injetadas com 14,8% de falsos positivos, qualificando-se como monitor de integridade das entradas em produção. Em conjunto, os três componentes delimitam papéis complementares: o MLP estima, o SOM valida a representação e a *Fuzzy* ART protege a fronteira do sistema.

Os resultados devem ser lidos dentro das limitações declaradas (rótulos derivados de perfis ocupacionais padronizados, e não de anúncios reais; rótulos prata transferidos por crosswalk e confinados ao treinamento; e baixa variância dos rótulos de Adaptabilidade). Os desdobramentos naturais seguem

dessas limitações: a anotação humana de uma amostra de vagas reais em português e inglês, permitindo ajuste fino do codificador e validação quantitativa no domínio de uso; e a investigação de rótulos com maior variância para adaptabilidade. No desenho do sistema, o módulo avaliado informa a aproximação entre candidatos e vagas, mas não decide sobre pessoas; as decisões permanecem ancoradas em evidência comportamental e revisão humana, condição que este trabalho reafirma como requisito, e não como concessão.

AGRADECIMENTOS

À professora Angela, pela disciplina DCE819 - Redes Neurais Artificiais (UNIFAL-MG), e à minha família, Cinara Silvia Alves da Rocha e Clibson Alves dos Santos, pelo apoio.

REFERÊNCIAS

Ali, M. M.; Qureshi, S. M.; Memon, M. S.; Mari, S. I.; Ramzan, M. B. Competency framework development for effective human resource management. SAGE Open, 11(2), 2021.

Baert, S.; Rotsaert, O.; Verhaest, D.; Omey, E. Skills, signals, and employability: an experimental investigation. European Economic Review, 123, 103374, 2020.

Bechtold, M.; Martínez-Herrera, E. et al. Skills or degree? The rise of skill-based hiring for AI and green jobs. Technological Forecasting and Social Change, 2025.

[AUTOR(ES) A CONFIRMAR]. Beyond degrees: analysis of skills-based versus credential-based hiring on employee performance and internal mobility. Management Journal of Research and Development, 7(1), 2026.

Carpenter, G. A.; Grossberg, S.; Rosen, D. B. Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system. Neural Networks, 4(6), 759-771, 1991.

Cybenko, G. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems, 2, 303-314, 1989.

Dipboye, R. L. Structured and unstructured selection interviews: beyond the job-fit model. Research in Personnel and Human Resources Management, 12, 79-123, 1994.

European Commission. ESCO: European Skills, Competences, Qualifications and Occupations. Publications Office of the European Union, 2024. Disponível em: https://esco.ec.europa.eu

He, K.; Zhang, X.; Ren, S.; Sun, J. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. Proceedings of the IEEE International Conference on Computer Vision (ICCV), 1026-1034, 2015.

Huffcutt, A. I.; Arthur, W., Jr. Hunter and Hunter (1984) revisited: interview validity for entry-level jobs. Journal of Applied Psychology, 79(2), 184-190, 1994.

Kohonen, T. The self-organizing map. Proceedings of the IEEE, 78(9), 1464-1480, 1990.

Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. International Conference on Learning Representations (ICLR), 2019.

McKinsey Global Institute. Human skills will matter more than ever in the age of AI. McKinsey & Company, 2025.

Moreno, A. L. Redes Neurais Artificiais: notas de aula da disciplina DCE819. Universidade Federal de Alfenas (UNIFAL-MG), 2026.

Nair, V.; Hinton, G. E. Rectified linear units improve restricted Boltzmann machines. Proceedings of the 27th International Conference on Machine Learning (ICML), 807-814, 2010.

OECD. Empowering the workforce in the context of a skills-first approach. OECD Skills Studies. OECD Publishing, Paris, 2025.

Reimers, N.; Gurevych, I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP), 3982-3992, 2019.

Rocha, J.; Kopp, B.; Zamarripa, N.; Sekhri, N. Widening the scope of job search: how human-centric skills and experience are redefining early careers. Manuscrito não publicado. Future Talent Council Emerging Leaders Ecosystem, 2026.

Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. Nature, 323, 533-536, 1986.

Sackett, P. R.; Zhang, C.; Berry, C. M.; Lievens, F. Resumes vs. application forms: why the stubborn reliance on resumes? Frontiers in Psychology, 13, 884205, 2022.

Sigelman, M.; Fuller, J. B.; Martin, A. Skills-based hiring: the long road from pronouncements to practice. Burning Glass Institute / Harvard Business School, 2024.

U.S. Department of Labor. O*NET Database. National Center for O*NET Development, 2026. Disponível em: https://www.onetcenter.org

Vedapradha, R. et al. Reimagining recruitment: traditional methods meet AI interventions, a 20-year assessment (2003-2023). Cogent Business & Management, 12(1), 2454319, 2025.

World Economic Forum. Future of jobs report 2025. World Economic Forum, Geneva, 2025.

EUROPEAN COMMISSION. Directorate-General for Employment, Social Affairs and Inclusion. The crosswalk between ESCO and O*NET: technical report. Brussels: European Commission, 2022. Disponível em: https://esco.ec.europa.eu/system/files/2022-12/ONET%20ESCO%20Technical%20Report.pdf.