Editions Ishango - UAC
Revue Internationale Multidisciplinaire Etincelle

Ed. Ishango

Article

CoBERT++ : recommandation contextuelle et explicable des profils d'employabilité des diplômés IT en Afrique subsaharienne

MUMBERE MOKYA Samy[1]*, Nsenge Mpia Héritier[2]

[1]Institut Supérieur de Commerce de Butembo (ISC/Butembo), Butembo, République Démocratique du Congo
[2] Université de l'Assomption au Congo, Butembo, République Démocratique du Congo

* Correspondant : mumberemokyasamy@gmail.com

Abstract: Cette étude présente CoBERT++, une extension contextuelle et explicable de CoBERT pour recommander des profils d'employabilité aux diplômés IT d'Afrique subsaharienne. Le corpus combine 355 observations réelles et 42 variables avec un enrichissement synthétique contrôlé, soit 6 942 enregistrements expérimentaux, auxquels sont associés douze indicateurs macroéconomiques, numériques et institutionnels. Dans le protocole local, CoBERT obtient MAP@5 = 0,431, nDCG@5 = 0,672 et ILD = 0,63. L'ajout du contexte externe porte MAP@5 à 0,492 (+14,2 %) et nDCG@5 à 0,718 (+6,8 %). Le réordonnancement MMR augmente l'ILD à 0,72, avec une perte relative de nDCG@5 annoncée comme inférieure ou égale à 3 %. Une validation technique limitée à la branche numérique et à 500 cas synthétiques classe S5, le volume de tweets et S2 en tête. LIME atteint R^2 = 0,853 et SHAP est stable (ϱ = 0,959 ; Jaccard Top-3 = 0,910), malgré un faible accord global SHAP–LIME (ϱ = −0,115). CoBERT++ améliore donc la pertinence et la diversité, mais requiert une validation réelle, inter-pays, équitable et centrée utilisateur.

Mot clés: CoBERT ; recommandation contextuelle ; employabilité ; diplômés IT ; Afrique subsaharienne ; MMR ; SHAP ; LIME.

Citation: Mumbere, M.S. & Mpia, H.N. CoBERT++ : recommandation contextuelle et explicable des profils d'employabilité des diplômés IT en Afrique subsaharienne. *Etincelle*. 2026, Vol. 26, no. 2. https://doi.org/10.61532/rime262143
Reçu : 25 juin 2026
Accepté : 29 juillet 2026
Publié : 04 août 2026

Note de l'éditeur: Ishango-uac reste neutre en ce qui concerne les revendications juridictionnelles dans les cartes géographiques publiées et les affiliations institutionnelles des auteurs.

1. Introduction

1.1. Contexte et problématique

L'insertion professionnelle des diplômés en technologies de l'information ne dépend pas uniquement de la maîtrise des outils numériques. Dans plusieurs pays d'Afrique subsaharienne, la guerre, la localisation de l'établissement de formation, les réseaux familiaux, la structure du marché du travail, l'accès à Internet et la stabilité institutionnelle influencent simultanément les possibilités d'emploi. Une recommandation fondée uniquement sur la proximité sémantique entre un profil étudiant et un métier risque donc de négliger les conditions concrètes dans lesquelles le diplômé devra chercher, obtenir et conserver un emploi. Cette difficulté est particulièrement importante dans les contextes où les données historiques sont rares, les trajectoires professionnelles peu documentées et les économies fortement exposées aux chocs politiques ou macroéconomiques (International Labour Organization, 2022 ; Mpia et al., 2022, 2023 ; Dake, 2025).

Les systèmes de recommandation ont largement démontré leur capacité à réduire la surcharge informationnelle et à personnaliser les choix, notamment dans le commerce électronique,

l'apprentissage en ligne et la recherche d'emploi (Jannach et al., 2011). Leur application à l'orientation universitaire soulève toutefois une exigence supplémentaire : la recommandation doit être pertinente, diversifiée et compréhensible. Une liste très précise mais composée de profils presque identiques limite l'exploration des possibilités ; inversement, une liste diversifiée mais peu pertinente peut désorienter l'utilisateur. De plus, les établissements et les étudiants doivent pouvoir comprendre les facteurs qui ont influencé le classement, surtout lorsque les données utilisées décrivent des caractéristiques sociales ou institutionnelles sensibles.

Le modèle CoBERT proposé par Mpia et al. (2023) répond à une partie de ce problème en combinant BERT avec des facteurs contextuels issus d'une enquête sur les diplômés IT en République Démocratique du Congo. Il recommande des profils d'employabilité à partir de la similarité entre la description d'un utilisateur et les profils de diplômés employés. Son apport principal est de montrer que l'employabilité doit être analysée avec des variables telles que le contexte sociopolitique, les compétences académiques et les relations entre diplômés et employeurs, en continuité avec les facteurs empiriquement étudiés dans le contexte congolais (Mpia et al., 2022). Néanmoins, le modèle initial ne représente pas explicitement l'évolution du marché du travail, l'environnement numérique ou la situation macroéconomique du pays au moment de la recommandation.

1.2. Question, objectif et contributions

La question étudiée est la suivante : dans quelle mesure l'intégration d'indicateurs externes, d'un mécanisme de diversification et d'une couche d'explicabilité améliore-t-elle la recommandation de profils d'employabilité destinée aux diplômés IT d'Afrique subsaharienne ? L'objectif est de concevoir et d'évaluer CoBERT++, une extension de CoBERT qui fusionne les variables individuelles avec des signaux macroéconomiques, numériques et institutionnels, puis applique Maximal Marginal Relevance (MMR) et des explications post-hoc par SHAP et LIME.

L'article apporte quatre contributions circonscrites aux éléments effectivement disponibles dans le mémoire. Premièrement, il décrit une architecture multimodale qui combine les représentations textuelles BERT/mBERT et douze indicateurs externes. Deuxièmement, il compare une reproduction locale de CoBERT avec l'extension contextuelle au moyen de MAP@5 et nDCG@5. Troisièmement, il mesure le compromis pertinence-diversité après réordonnancement MMR au moyen de l'Intra-List Diversity. Quatrièmement, il rapporte une validation technique SHAP/LIME réalisée sur la branche numérique sauvegardée. Cette dernière ne doit pas être interprétée comme une validation empirique sur les diplômés réels, car le corpus textuel exact, le tokenizer et la définition des cinq facteurs statiques ne sont pas disponibles dans l'archive technique.

2. Cadre conceptuel et travaux connexes

2.1. Recommandation sensible au contexte

Un système de recommandation sensible au contexte considère que la pertinence d'un item dépend non seulement de l'utilisateur et de l'item, mais aussi des conditions de décision. Adomavicius et Tuzhilin (2011) distinguent notamment l'intégration du contexte avant, pendant ou après le calcul du score. Dans le cas de l'employabilité, le contexte peut inclure le pays, l'année, la dynamique économique, le chômage, la disponibilité des infrastructures numériques et la stabilité des institutions. Ces dimensions ne remplacent pas les compétences individuelles ; elles permettent de moduler leur valeur selon les opportunités réellement accessibles. Une revue systématique récente confirme que la représentation du contexte, son mode d'intégration et la reproductibilité expérimentale restent des enjeux centraux des systèmes de recommandation sensibles au contexte (Mateos & Bellogín, 2025).

CoBERT adopte un filtrage basé sur le contenu centré sur les profils. BERT transforme les descriptions textuelles en représentations sémantiques et le système identifie les profils d'employabilité les plus proches (Devlin et al., 2019 ; Mpia et al., 2023). Cette approche limite le problème de démarrage à froid, car elle peut fonctionner sans historique détaillé d'interactions. Elle reste cependant sensible à la qualité de la description, aux biais du corpus préentraîné et à la

disponibilité de variables contextuelles pertinentes. CoBERT++ prolonge cette logique en ajoutant un vecteur externe associé au pays et à l'année.

2.2. Pertinence, diversité et explicabilité

La qualité d'un classement est évaluée par des métriques qui tiennent compte à la fois de la présence des recommandations pertinentes et de leur position. Le MAP@k résume la précision moyenne dans les premiers rangs, tandis que le nDCG@k accorde davantage de poids aux éléments pertinents placés en tête de liste (Järvelin & Kekäläinen, 2002). Ces métriques ne mesurent pas directement la redondance. MMR réordonne les candidats en combinant pertinence et dissimilarité, ce qui permet d'augmenter la diversité intra-liste sans reconstruire entièrement le modèle (Carbonell & Goldstein, 1998).

L'explicabilité répond à une autre dimension de la qualité. SHAP estime la contribution de chaque variable à partir des valeurs de Shapley et peut être utilisé pour une lecture globale ou locale du modèle (Lundberg & Lee, 2017). LIME construit un modèle interprétable dans le voisinage d'une prédiction et fournit des coefficients locaux (Ribeiro et al., 2016). Les deux méthodes sont complémentaires, mais elles ne sont pas nécessairement concordantes : SHAP dépend notamment du jeu de référence et LIME de la définition du voisinage, du nombre de perturbations et de la stabilité de l'approximation. Une évaluation XAI rigoureuse doit donc documenter la fidélité, la stabilité et l'accord entre méthodes, conformément aux cadres d'évaluation quantitative de l'IA explicable et de la recommandation explicable (Nauta et al., 2023 ; Zhang & Chen, 2020).

3. Méthodologie

3.1. Données et variables

La source micro-individuelle est l'enquête CoBERT déposée dans Harvard Dataverse (Mpia, 2022). Le corpus réel comprend 355 répondants répartis dans les quatre zones linguistiques de la RDC et 42 variables décrivant les dimensions démographiques, académiques, socioéconomiques et contextuelles. La variable cible correspond au statut d'emploi. Parmi les variables utilisées dans le modèle initial figurent la guerre dans la zone de naissance ou d'étude, la localisation de l'institution, la compétence dans la spécialisation IT, la compétence IT générale, la situation financière de la famille et la présence de proches dans l'administration, la politique ou les entreprises.

Afin d'étudier une architecture plus complexe malgré la taille limitée de l'échantillon, le mémoire applique un enrichissement synthétique contrôlé par CTGAN. Après les contrôles de cohérence, le corpus expérimental final comporte 6 942 enregistrements. Cette augmentation facilite les expérimentations, mais elle ne crée pas une nouvelle preuve de terrain : les résultats doivent rester rattachés au socle initial de 355 observations. CTGAN est utilisé comme outil de simulation tabulaire et non comme substitut à une collecte réelle plus large (Xu et al., 2019).

Douze indicateurs externes sont associés aux profils par pays et par année : PIB par habitant en parité de pouvoir d'achat, chômage des jeunes, inflation, indice de Gini, pénétration d'Internet, dépenses publiques d'éducation, indice de développement humain, indice global de paix, indice de perception de la corruption, taux de vacance d'emploi, population totale et chômage national. Les sources prévues comprennent les bases WDI de la Banque mondiale, ILOSTAT, l'UIT, le PNUD, l'Institute for Economics and Peace et Transparency International. Le protocole couvre sept pays entre 2018 et 2024, mais les éléments disponibles ne fournissent des données pilotes que pour 60 profils du Nigeria, du Cameroun et du Soudan du Sud.

Tableau 1 : Données, composantes et rôle dans CoBERT++

Composante	Contenu disponible	Rôle dans le modèle
Enquête CoBERT	355 observations réelles, 42 variables, quatre zones linguistiques	Profil individuel et cible d'emploi
Corpus expérimental	6 942 enregistrements après enrichissement CTGAN et contrôles	Entraînement et comparaison locale
Contexte externe	12 indicateurs économiques, numériques et institutionnels	Modulation du score selon le pays et l'année
Diversification	Réordonnancement MMR	Réduction de la redondance des profils proposés
Explicabilité	SHAP global/local et LIME local	Justification technique des scores

Composante	Contenu disponible	Rôle dans le modèle
Validation géographique	Pilote de 60 profils dans trois pays	Indication de transférabilité, sans généralisation statistique

3.2. Architecture CoBERT++

Le pipeline proposé comporte cinq étapes. Premièrement, les variables individuelles sont transformées en une description textuelle structurée, puis encodées par BERT ou mBERT. Deuxièmement, les douze indicateurs externes sont nettoyés, normalisés et projetés dans un espace numérique de dimension réduite. Troisièmement, les représentations textuelle et numérique sont fusionnées afin de calculer un score d'adéquation pour chaque profil d'employabilité. Quatrièmement, MMR réordonne les candidats en pénalisant les profils trop similaires. Cinquièmement, SHAP et LIME expliquent le score sans modifier le classement. Cette séparation entre prédiction, diversification et explication permet d'identifier la contribution propre de chaque module.

Le mémoire décrit une implémentation Python utilisant HuggingFace Transformers pour l'encodeur, Flask pour l'interface et des bibliothèques XAI pour les explications. Toutefois, la reproductibilité complète du modèle textuel demeure limitée : l'archive technique contient les poids sauvegardés, mais pas le tokenizer, le corpus textuel exact ni le dictionnaire des facteurs statiques S1 à S5. En conséquence, l'évaluation XAI rapportée ci-dessous maintient la représentation textuelle constante et étudie uniquement la branche numérique. Cette décision évite d'attribuer aux explications une portée que les artefacts disponibles ne permettent pas de démontrer.

3.3. Variantes, protocole et métriques

Quatre variantes principales sont distinguées. M0 reproduit CoBERT avec les variables individuelles et constitue la baseline locale. M1 ajoute les douze indicateurs externes. M2 applique MMR aux scores de M1 pour améliorer la diversité. M3 ajoute les explications SHAP et LIME ; comme l'explication est post-hoc, les métriques de classement ne sont pas recalculées. Une cinquième série d'essais examine BERT, mBERT et XLM-RoBERTa ainsi que la transférabilité géographique, mais aucun score détaillé par encodeur ou par pays n'est disponible dans le mémoire.

Les comparaisons principales sont effectuées à $k = 5$. MAP@5 mesure la capacité à placer les profils pertinents dans les cinq premiers résultats et nDCG@5 évalue la qualité de leur ordre. L'ILD quantifie la diversité moyenne entre les éléments d'une même liste. Pour l'explicabilité, le protocole retient la moyenne des valeurs absolues SHAP, le R^2 du modèle local LIME, la variation de score après ablation, la corrélation de Spearman sous perturbations et le Jaccard du Top-3. Les valeurs publiées dans l'article CoBERT à Top-N = 4 sont conservées comme repère externe, sans être comparées directement aux résultats locaux à $k = 5$.

3.4. Protocole technique SHAP/LIME

L'analyse XAI utilise le fichier de poids sauvegardé du prototype. Un plan Latin Hypercube reproductible de 500 observations normalisées dans l'intervalle [0,1] est généré, conformément au domaine numérique du script d'entraînement. Trente observations servent de fond SHAP et cinquante cas sont expliqués. LIME est appliqué au même modèle et au même espace de variables. La fidélité est testée en retirant simultanément les trois facteurs les plus importants et en comparant la variation obtenue à l'ablation de trois variables aléatoires. La stabilité est mesurée sur dix profils, chacun soumis à cinq perturbations gaussiennes de paramètre $\sigma = 0{,}02$. L'accord entre SHAP et LIME est résumé par la corrélation de Spearman des importances et par le recouvrement du Top-3.

4. Résultats

4.1. Pertinence et diversité

L'article CoBERT de 2023 rapporte un nDCG de 0,99 à Top-N = 4 et un MAP de validation de 0,50. Dans le protocole local harmonisé à $k = 5$, M0 obtient un MAP@5 de 0,431, un nDCG@5 de 0,672 et un ILD de 0,63. L'écart avec la publication initiale ne constitue pas une baisse directement interprétable, car la profondeur de classement et les conditions expérimentales diffèrent. La comparaison pertinente est celle de M0 et M1 dans le même protocole local.

L'ajout du contexte externe améliore les deux métriques de pertinence. Le MAP@5 passe de 0,431 à 0,492, soit un gain relatif de 14,2 %. Le nDCG@5 augmente de 0,672 à 0,718, soit 6,8 %. Le gain plus élevé du MAP suggère que les indicateurs externes aident surtout à identifier les profils pertinents parmi les premiers candidats, tandis que leur effet sur l'ordonnancement fin

est plus modéré. Ces résultats soutiennent l'hypothèse selon laquelle les opportunités d'emploi doivent être interprétées à la lumière de la situation économique et institutionnelle.

MMR porte l'ILD de 0,63 à 0,72, soit une progression de 14,3 %. Le mémoire indique une perte relative de nDCG@5 inférieure ou égale à 3 %. En appliquant cette borne au score de M1, le nDCG@5 après réordonnancement se situe entre 0,696 et 0,718. Même dans le scénario inférieur, il demeure supérieur à la baseline locale de 0,672. Le MAP@5 après MMR n'a pas été recalculé séparément ; la valeur de 0,492 correspond donc au modèle M1 avant réordonnancement.

Tableau 2 : Résultats comparatifs des variantes de CoBERT++

Variante	MAP@5	nDCG@5	ILD	Interprétation
M0 – CoBERT reproduit	0,431	0,672	0,63	Baseline locale
M1 – + contexte externe	0,492	0,718	0,63	Pertinence accrue : +14,2 % MAP ; +6,8 % nDCG
M2 – + MMR	0,492*	0,696–0,718**	0,72	Diversité accrue : +14,3 % ILD
M3 – + XAI	—	—	—	Explication post-hoc ; classement non recalculé

* Le MAP@5 après MMR n'a pas été recalculé séparément.
** Intervalle dérivé d'une perte relative maximale de 3 % du nDCG@5 de M1.

4.2. Résultats de l'explicabilité

L'importance globale SHAP place le facteur statique S5 au premier rang avec une moyenne absolue de 0,0822 point de score sur 100, suivi du volume de tweets (0,0791) et du facteur statique S2 (0,0722). Cette hiérarchie décrit le comportement de la branche numérique sauvegardée. Elle ne permet pas d'interpréter substantiellement S2 ou S5, car leur signification n'est pas documentée. Les contributions ne doivent pas non plus être généralisées au corpus réel, puisque les explications ont été calculées sur des observations synthétiques normalisées.

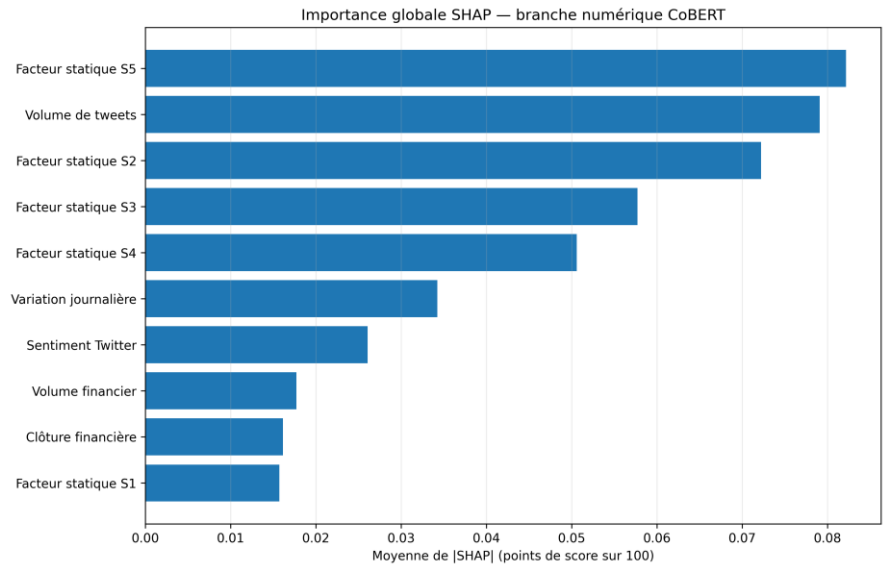

Figure 1 : Importance globale SHAP de la branche numérique sauvegardée du prototype CoBERT.

Pour le profil synthétique retenu comme exemple, le score prédit est de 50,47/100. LIME obtient un R^2 local de 0,853 et un R^2 moyen de 0,857 sur les cas évalués, ce qui indique que le modèle linéaire substitut reproduit correctement le voisinage étudié. La fidélité par ablation confirme l'utilité des facteurs dominants : retirer le Top-3 modifie le score de 0,1288 en moyenne, avec un intervalle de 0,1058 à 0,1542, contre 0,0859 pour trois variables aléatoires. Le ratio de 1,50 signifie que les variables classées par SHAP ont un effet plus marqué que le contrôle aléatoire.

Les explications SHAP restent stables sous de faibles perturbations. La corrélation de Spearman moyenne des rangs atteint 0,959, avec un intervalle de 0,948 à 0,969, et le Jaccard du Top-3 atteint 0,910, avec un intervalle de 0,850 à 0,960. En revanche, l'accord global des importances SHAP et LIME est faible et négatif ($\varrho = -0{,}115$), malgré un recouvrement de 66,7 % du Top-3. Les deux méthodes identifient donc certains facteurs communs, mais elles ne décrivent pas de manière identique la structure globale du modèle.

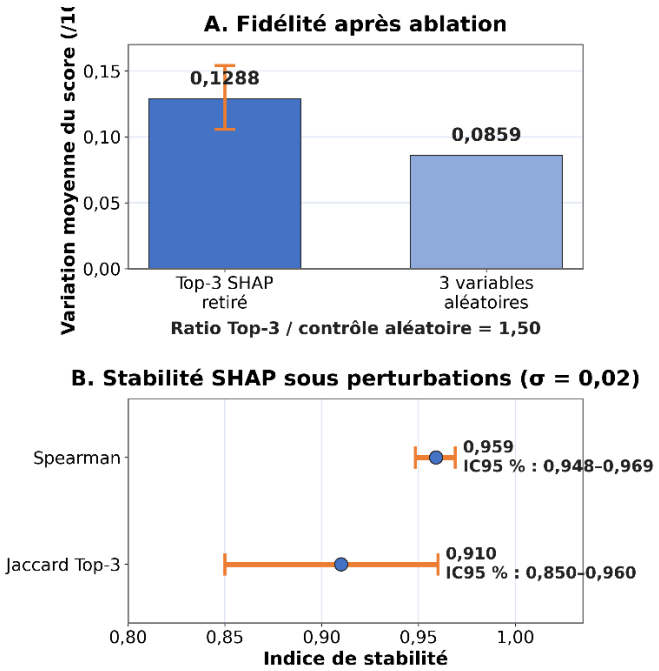

Figure 2 : Fidélité par ablation et stabilité des explications SHAP sous perturbations.

Tableau 3 : Synthèse de l'évaluation technique SHAP/LIME

Dimension	Résultat calculé	Interprétation prudente
Importance SHAP	Top-3 : S5 = 0,0822 ; tweets = 0,0791 ; S2 = 0,0722	Hiérarchie technique sur 50 cas synthétiques
Fidélité LIME	R^2 local = 0,853 ; R^2 moyen = 0,857	Bonne approximation dans le voisinage étudié
Ablation	Δ Top-3 = 0,1288 ; contrôle = 0,0859 ; ratio = 1,50	Facteurs dominants plus influents que le contrôle
Stabilité SHAP	Spearman = 0,959 ; Jaccard Top-3 = 0,910	Stabilité élevée sous bruit faible
Accord SHAP–LIME	Spearman = −0,115 ; Top-3 commun = 66,7 %	Méthodes complémentaires, non interchangeables

4.3. Robustesse linguistique et géographique

Le mémoire prévoit une comparaison de BERT, mBERT et XLM-RoBERTa ainsi qu'une validation dans plusieurs pays. Les artefacts disponibles ne fournissent cependant ni scores détaillés par encodeur ni valeurs MAP/nDCG par pays. Le pilote de 60 profils provenant du Nigeria, du Cameroun et du Soudan du Sud indique que le pipeline peut être appliqué à des contextes extérieurs à la RDC, mais il ne permet pas d'établir une généralisation statistique. La transférabilité doit donc être qualifiée de partielle et exploratoire.

5. Discussion

5.1. Apport du contexte externe

L'amélioration conjointe du MAP@5 et du nDCG@5 montre que les variables externes ajoutent une information utile aux caractéristiques individuelles. Un profil technique identique ne présente pas la même probabilité d'insertion lorsque le chômage des jeunes, l'accès à Internet, l'inflation ou la stabilité institutionnelle diffèrent. La fusion contextuelle permet au système de rapprocher la recommandation de la situation réelle du marché. Elle ne démontre toutefois pas une relation causale entre chaque indicateur et l'emploi : les résultats mesurent une amélioration prédictive dans le protocole expérimental, pas l'effet économique isolé de chaque variable.

Le gain relatif de 14,2 % du MAP est particulièrement pertinent pour une application d'orientation. Il signifie que les profils utiles apparaissent plus fréquemment dans les premiers résultats. Le gain de 6,8 % du nDCG indique que l'ordre s'améliore aussi, mais dans une proportion plus faible. Ce contraste suggère que le contexte contribue davantage à la sélection des candidats pertinents qu'à leur hiérarchisation précise. Des analyses d'ablation par groupe d'indicateurs seraient nécessaires pour déterminer si cette amélioration provient principalement de l'économie, de l'infrastructure numérique ou de la gouvernance.

5.2. Compromis diversité–pertinence

L'augmentation de l'ILD de 0,63 à 0,72 confirme l'utilité du réordonnancement MMR. Une orientation professionnelle ne doit pas enfermer l'étudiant dans une série de variantes presque identiques d'un même profil. La diversité permet d'exposer des parcours alternatifs compatibles avec ses compétences. La perte maximale annoncée de 3 % du nDCG reste modérée et le score demeure supérieur à la baseline. Le compromis observé est donc favorable, tout en nécessitant une mesure explicite du MAP et de la couverture après réordonnancement dans une réplication complète.

5.3. Portée de l'explicabilité

La couche XAI ajoute une transparence technique au prototype. SHAP fournit une hiérarchie globale stable et LIME une explication locale fidèle dans le voisinage d'un profil. L'ablation renforce la crédibilité de l'importance SHAP, car les variables dominantes modifient davantage le score que des variables aléatoires. La faible corrélation globale entre SHAP et LIME rappelle toutefois qu'une explication n'est pas une vérité unique. Les méthodes répondent à des hypothèses différentes et doivent être présentées ensemble avec leurs paramètres, leur jeu de référence et leur domaine de validité.

La principale limite de cette validation est la nature des données explicatives. Le modèle textuel ne peut pas être reconstruit fidèlement sans le tokenizer et le corpus exacts. Les cinq facteurs statiques ne sont pas définis, ce qui empêche de traduire leurs contributions en recommandations compréhensibles pour un étudiant ou un conseiller. Avant un déploiement, l'interface devra remplacer les étiquettes S1–S5 par des noms documentés, afficher la direction et l'intensité des facteurs et préciser que les explications décrivent le modèle, non une causalité sociale.

5.4. Implications et limites

Pour les universités, CoBERT++ peut soutenir les services d'orientation, le suivi des diplômés et la révision des programmes. Le système pourrait signaler les compétences à renforcer et proposer plusieurs trajectoires professionnelles plutôt qu'un seul choix. Pour les décideurs publics, les résultats agrégés pourraient aider à identifier les contraintes contextuelles qui réduisent l'utilité de certaines formations. Dans les deux cas, le modèle doit rester un outil d'aide : la décision finale doit associer l'étudiant, le conseiller et les informations locales sur les opportunités.

Cinq limites restreignent la portée des résultats. Le socle réel de 355 répondants est modeste et l'enrichissement CTGAN ne remplace pas une nouvelle collecte. Les indicateurs annuels ne capturent pas toutes les disparités infranationales ni les changements rapides du marché. Le pilote inter-pays repose sur 60 profils sans scores détaillés. L'évaluation XAI est limitée à la branche numérique et à des données synthétiques. Enfin, aucune étude d'équité entre genres, régions ou niveaux socioéconomiques, ni aucune évaluation d'utilisabilité auprès des étudiants et conseillers, n'est rapportée. Ces limites interdisent de présenter CoBERT++ comme un système prêt pour une décision autonome.

6. Conclusion

CoBERT++ prolonge CoBERT en combinant les variables individuelles des diplômés IT avec douze indicateurs décrivant l'économie, le marché du travail, l'environnement numérique et les institutions. Dans le protocole local, cette fusion porte le MAP@5 de 0,431 à 0,492 et le nDCG@5 de 0,672 à 0,718. MMR augmente l'ILD de 0,63 à 0,72 tout en maintenant le nDCG@5 dans une plage compatible avec une perte relative maximale de 3 %. L'analyse SHAP/LIME montre qu'une explication globale et locale est techniquement possible, avec une bonne fidélité LIME et une forte stabilité SHAP.

Ces résultats soutiennent l'idée que l'orientation professionnelle doit tenir compte du contexte dans lequel les compétences seront mobilisées. Ils ne suffisent toutefois pas à démontrer une généralisation inter-pays ou une interprétation empirique complète des facteurs XAI. Les travaux futurs devront restaurer le pipeline textuel reproductible, documenter les facteurs S1–S5, collecter davantage de profils réels par pays, intégrer des offres d'emploi locales et actualisées, mesurer l'équité et conduire une évaluation utilisateur. CoBERT++ constitue ainsi une base méthodologique prometteuse pour une recommandation d'employabilité contextualisée, diversifiée et transparente, sous réserve d'une validation externe rigoureuse.

Contributions des auteurs : Conceptualisation, M.M.S. et N.M.H. ; méthodologie, M.M.S. et N.M.H. ; analyse et rédaction initiale, M.M.S. ; supervision, validation et révision critique, N.M.H. Les deux auteurs ont approuvé la version finale du manuscrit.

Sponsor financier : Cette recherche n'a reçu aucun financement spécifique.

Disponibilité des données : Le jeu de données CoBERT est référencé dans Harvard Dataverse sous le DOI 10.7910/DVN/SI2RP1. Les résultats XAI rapportés proviennent de l'évaluation technique décrite dans le mémoire ; les artefacts textuels complets nécessaires à une reproduction intégrale ne sont pas tous disponibles.

Remerciements : Les auteurs remercient les institutions et les répondants ayant contribué à l'enquête CoBERT.

Conflits d'intérêt : Les auteurs déclarent ne connaître aucun conflit d'intérêt relatif à cette étude.

Références

1. Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. Dans F. Ricci, L. Rokach, B. Shapira, & P. B. Kantor (dir.), Recommender systems handbook (pp. 217–253). Springer. https://doi.org/10.1007/978-0-387-85820-3_7
2. Carbonell, J., & Goldstein, J. (1998). The use of MMR, diversity-based reranking for reordering documents and producing summaries. Dans Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 335–336). Association for Computing Machinery. https://doi.org/10.1145/290941.291025
3. Dake, D. K. (2025). Information technology graduates employability prediction model in a low-income country using tree-based machine learning classifiers. Discover Global Society, 3, article 158. https://doi.org/10.1007/s44282-025-00304-3
4. Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. Dans Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (vol. 1, pp. 4171–4186). Association for Computational Linguistics. https://doi.org/10.18653/v1/N19-1423
5. International Labour Organization. (2022). Global employment trends for youth 2022: Investing in transforming futures for young people. https://doi.org/10.54394/QSMU1809
6. Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). Recommender systems: An introduction. Cambridge University Press. https://doi.org/10.1017/CBO9780511763113
7. Järvelin, K., & Kekäläinen, J. (2002). Cumulated gain-based evaluation of information retrieval techniques. ACM Transactions on Information Systems, 20(4), 422–446. https://doi.org/10.1145/582415.582418
8. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30. https://papers.nips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html
9. Mateos, P., & Bellogín, A. (2025). A systematic literature review of recent advances on context-aware recommender systems. Artificial Intelligence Review, 58, article 20. https://doi.org/10.1007/s10462-024-10939-4
10. Mpia, H. N. (2022). Replication data for: Congolese IT graduates employability data [Jeu de données]. Harvard Dataverse. https://doi.org/10.7910/DVN/SI2RP1
11. Mpia, H. N., Mwendia, S. N., & Mburu, L. W. (2022). Predicting employability of Congolese information technology graduates using contextual factors: Towards sustainable employability. Sustainability, 14(20), 13001. https://doi.org/10.3390/su142013001
12. Mpia, H. N., Mburu, L. W., & Mwendia, S. N. (2023). CoBERT: A contextual BERT model for recommending employability profiles of information technology students in unstable developing countries. Engineering Applications of Artificial Intelligence, 125, 106728. https://doi.org/10.1016/j.engappai.2023.106728
13. Nauta, M., Trienes, J., Pathak, S., Nguyen, E., Peters, M., Schmitt, Y., Schlötterer, J., van Keulen, M., & Seifert, C. (2023). From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable AI. ACM Computing Surveys, 55(13s), article 295. https://doi.org/10.1145/3583558
14. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. Dans Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939778
15. Xu, L., Skoularidou, M., Cuesta-Infante, A., & Veeramachaneni, K. (2019). Modeling tabular data using conditional GAN. Advances in Neural Information Processing Systems, 32. https://papers.nips.cc/paper/2019/hash/254ed7d2de3b23ab10936522dd547b78-Abstract.html
16. Zhang, Y., & Chen, X. (2020). Explainable recommendation: A survey and new perspectives. Foundations and Trends in Information Retrieval, 14(1), 1–101. https://doi.org/10.1561/1500000066