

Exprémer continguts educatius amb models de llenguatge extens: com deia Jack l'Esbudellador, és una qüestió de trossejar bé

Daniel Amo-Filva
Enginyeria

La Salle Campus Barcelona, Universitat Ramon Llull
Barcelona, Espanya
daniel.amo@salle.url.edu

Resum— Aquest article discuteix la importància del context en la diferenciació entre humans i màquines, destacant com el context cultural, social, econòmic i polític contribueix a la nostra humanitat. Aquest treball explora mètodes per apropar les màquines a l'experiència humana proporcionant-los un context humà i facilitant així una relació ètica amb elles. A més, s'aborden les limitacions dels LLMs, com les al·lucinacions, i es proposen tècniques per a mitigar-les i millorar la seva precisió i aplicabilitat, especialment en entorns educatius. El treball es centra en la utilització d'aquests models com a cas d'ús en l'assignatura Projectes Web I de La Salle Campus Barcelona, Universitat Ramon Llull. Es descriu un procés per contextualitzar continguts educatius per a la seva utilització en LLMs, incloent tècniques per convertir textos en contextos útils per aquests models.

Paraules clau—educació, models de llenguatge extens, context, immersió, indicació, RAG

I. INTRODUCCIÓ

Les màquines computacionals amb llampades d'intel·ligència artificial general (IAG) [1] es diferencien dels humans per molts aspectes. El context és el tret fonamental diferenciador. Hem referenciat a context com a tot allò que engloba cultures, societats, economies i polítiques humanes. El context ens fa humans, a les màquines les fa les altres, novingudes, foranes i desconegudes. Les màquines es converteixen en el rostre diferent de Lévinas: “En l'expressió un ser es presenta en sí mateix” [2] (p. 213). És potser per això que passem les màquines per un procés d'homomorfisme, per apropar-les a nosaltres en una relació d'ètica per a màquines positiva per ambdues parts. Per aconseguir IA, i generativa, hem generat models de llenguatge natural basats en algorismes matemàtics que emulen neurones artificials i els hi hem dotat de context humà creant petites simulacions de consciències, unes més properes que d'altres a les característiques del pensament humà. Aquesta fita ens permet relacionar-nos amb les màquines, comunicar-nos-hi amb llenguatge natural i dotar-les d'expressió, convertint-les en el rostre de Lévinas (llegir primer paràgraf).

A. Arreglant les falles del sistema

Emperò, els LLM no són perfectes, i a vegades surten del context real, segueixen un argumentari dotat d'irrealitats i, per tant, cal acotar-les en la seva aplicació en entorns de producció. Yao et al. [4] ho diuen clarament: “encara no podem confiar completament en la seva resposta, ja que els LLM pateixen al·lucinacions: fabriquen fets inexistents”. Una manera de mitigar els efectes de les al·lucinacions en els LLM és usar tècniques com la Injecció de Coneixement (Knowledge Injection, KI). En aquest sentit, Martino et al. [5] inclouen text rellevant provinent de grafs de coneixement en les indicacions d'LLM; d'aquesta manera milloren les respostes i redueixen les al·lucinacions dels LLMs.

B. Un altre mètode: Retrieval-Augmented Generation (RAG)

En clau universitària, un exemple d'ús d'LLM a educació és interactuar amb els continguts d'assignatures, de manera que, per exemple, el docent pot millorar i crear noves activitats d'aprenentatge, i els alumnes poden autoavaluar-se o entendre els continguts des de diferents perspectives. En aquest treball faig una proposta per a mitigar al·lucinacions i interactuar amb continguts educatius mitjançant LLMs. Segueixo el mètode KI per a mitigar al·lucinacions, però usant la tècnica RAG [6], [7] en comptes de grafs de coneixement. La tècnica RAG consisteix en usar models de llenguatge extens pre-entrenats junt amb un mecanisme diferenciable d'accés a memòria no paramètrica, normalment com a base de dades de vectors. Això permet que el model presenti un seguit de capacitats per omisió que es veuen augmentades amb coneixement extern i s'aconsegueixin resultats actualitzats. recuperar context extern. Addicionalment, proposo maneres d'estructurar el text dels continguts segons tasques a executar i tècniques RAG segons l'estructura dels textos dels continguts.

II. CAS D'ÚS

El cas d'ús s'emmarca en l'assignatura Projectes Web I de grau d'enginyeries de La Salle Campus Barcelona, Universitat Ramon Llull. L'objectiu és fer disponible una eina que faciliti

l'extracció de context a partir de continguts d'una assignatura per poder incloure'l en la indicació a fer a qualsevol LLM. Com a contingut per demostrar el procés de contextualització s'utilitza l'apartat "setup()" del capítol "Composition API" de la documentació de Vue, un framework de desenvolupament web front-end, i en format Markdown. D'aquesta manera tant docents com alumnes poden interactuar amb els continguts textuais fent-los servir dins la indicació.

III. EXPERIMENTS I RESULTATS

Per a poder convertir contingut textual en context per LLMs s'ha de realitzar el següent pipeline: 1) Trossejar. Partir el text en trossets més petits a partir de regles o models de tokenitzat, 2) Immersió. Convertir cada tros de text resultant a vectors computables, 3) Emmagatzemar. Guardar els vectors resultants a una base de dades de vectors, 4) Cercar. Cercar per similitud aquells vectors que s'assemblen a la tasca a executar per l'LLM. El procés de trossejar, àmbit d'estudi d'aquest treball, és el més delicat. En aquest treball he configurat diferents trossejadors, en concret els disponibles dins la llibreria LangChain (veure Figura 1), per a seleccionar el més adequat considerant les tasques:

- Extreure conclusions principals: l'histograma hauria de mostrar una concentració de llargada moderada a més llarga. El trossejador captura frases completes (pensaments incomplets) i inclús paràgrafs sencers (molt detall) que puguin contenir les conclusions principals.
- Resumir contingut en punts principals: l'histograma ha de tenir una forma equilibrada amb un pic a les longituds que normalment representen frases completes però concises. Això indica que el trossejador captura segments que probablement siguin punts coherents, adequats per a resums. La distribució no s'ha d'orientar massa cap a molta llargada per evitar la verbositat, o cap a frases curtes que poden perdre el context.
- Generar exercicis d'avaluació de diferents parts del contingut: l'histograma ha de mostrar una distribució més uniforme en un rang de llargades, reflectint una barreja de longituds de frases i paràgrafs. Això suggeriria que el trossejador pot capturar tant detalls específics (extensió curta a mitjana) com conceptes o context més amplis (extensió mitjana a llarga), que són necessaris per elaborar preguntes d'avaluació exhaustives que cobreixen diferents aspectes del contingut.

En aquest cas d'ús, una longitud de 200 és adient per a les tres tasques, on el trossejador RecursiveCharacterSplitter és més adequat considerant que forma una gaussiana que abraça frases i paràgrafs.

CONCLUSIONS

En aquest estudi, es mostra com l'ús de models de llenguatge extens pot ser extremadament beneficiós en contextos educatius, especialment quan es proporciona un context adequat. La implementació de tècniques com la Injecció de Coneixement i la Retrieval-Augmented Generation pot augmentar significativament la precisió i la funcionalitat dels LLMs, mitgant així problemes com les al·lucinacions. A través del cas

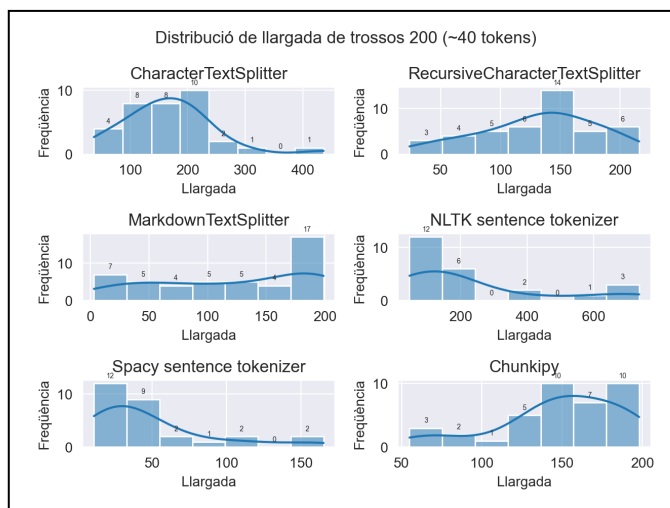


Fig. 1 Resultat dels diferents mètodes de trossejar de LangChain

d'ús a l'assignatura Projectes Web I de La Salle Campus Barcelona, Universitat Ramon Llull, he evidenciat com aquests models poden ser utilitzats per enriquir l'experiència educativa, tant per a professors com per a estudiants. Aquesta recerca obre noves possibilitats per a investigacions futures en l'àmbit de la interacció entre LLMs i usuaris humans en diferents àmbits, subratllant la importància de proporcionar un context rellevant i actualitzat. La contextualització apropiada dels LLMs serà clau per a la seva aplicació efectiva i ètica en la societat.

REFERENCES

- [1] S. Bubeck *et al.*, "Sparks of Artificial General Intelligence: Early experiments with GPT-4." arXiv, Apr. 13, 2023. doi: 10.48550/arXiv.2303.12712.
- [2] E. Lévinas, "Totalidad e infinito: Ensayo sobre la exterioridad," in *Totalidad e infinito: Ensayo sobre la exterioridad*, Salamanca: Sígueme, 1999, p. 315. Accessed: Jan. 05, 2024. [Online]. Available: <https://pesquisa.bvsalud.org/portal/resource/pt/psa-52091>
- [3] L. Gao *et al.*, "The Pile: An 800GB Dataset of Diverse Text for Language Modeling." arXiv, Dec. 31, 2020. doi: 10.48550/arXiv.2101.00027.
- [4] J.-Y. Yao, K.-P. Ning, Z.-H. Liu, M.-N. Ning, and L. Yuan, "LLM Lies: Hallucinations are not Bugs, but Features as Adversarial Examples." arXiv, Oct. 04, 2023. doi: 10.48550/arXiv.2310.01469.
- [5] A. Martino, M. Iannelli, and C. Truong, "Knowledge Injection to Counter Large Language Model (LLM) Hallucination," in *The Semantic Web: ESWC 2023 Satellite Events*, C. Pesquita, H. Skaf-Molli, V. Efthymiou, S. Kirrane, A. Ngonga, D. Collarana, R. Cerqueira, M. Alam, C. Trojahn, and S. Hertling, Eds., in Lecture Notes in Computer Science. Cham: Springer Nature Switzerland, 2023, pp. 182–185. doi: 10.1007/978-3-031-43458-7_34.
- [6] P. Lewis *et al.*, "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020, pp. 9459–9474. Accessed: Jan. 05, 2024. [Online]. Available: <https://proceedings.neurips.cc/paper/2020/hash/6b493230205f780e1bc26945df7481e5-Abstract.html>
- [7] W. Yu, "Retrieval-augmented Generation across Heterogeneous Knowledge," in *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Student Research Workshop*, D. Ippolito, L. H. Li, M. L. Pacheco, D. Chen, and N. Xue, Eds., Hybrid: Seattle, Washington + Online: Association for Computational Linguistics, Jul. 2022, pp. 52–58. doi: 10.18653/v1/2022.naacl-srw.7.