

Padrões de Interação em Tutoria por IA: Perfis Conversacionais Compatíveis com Aprendizagem Autorregulada

Otacilio Saraiva Maia Neto¹, Rafael Ferreira Leite de Mello^{1,2},
Taciana Pontual Falcão², André Câmara Alves do Nascimento^{1,2}

¹CESAR School

osmn@cesar.school

²Universidade Federal Rural de Pernambuco

{rafael.mello,taciana.pontual,andre.camara}@ufrpe.br

Abstract. *This paper investigates whether interaction traces from AI tutoring dialogues can reveal interpretable behavioral profiles compatible with self-regulated learning. We analyze the public StudyChat dataset, composed of 2,214 conversations from 203 students across seven programming assignments in an undergraduate artificial intelligence course. Student turns were mapped to dialogue acts and modeled as first- and second-order Markov transition matrices at the student-assignment level. After vectorization, standardization, UMAP dimensionality reduction, and K-means clustering, the resulting groups were interpreted using structural metrics and Heuristic Miner process models. The second-order representation yielded the most coherent solution, with six interaction profiles ranging from rigid directive behavior to dense exploratory patterns. A longitudinal analysis based on masked Hamming distance showed that many students moved between profiles across assignments rather than remaining in a single stable mode. The main contribution is an explainable learning analytics pipeline for conversational educational data, showing how dialogue-based traces can support cautious theory-informed interpretations relevant to adaptive AI tutors.*

Keywords: learning analytics, self-regulated learning, AI tutoring, dialogue acts, process mining.

Resumo. *Este artigo investiga se rastros interacionais de diálogos com tutores baseados em Inteligência Artificial podem revelar perfis comportamentais interpretáveis compatíveis com processos de aprendizagem autorregulada. Analisamos o conjunto de dados público StudyChat, composto por 2.214 conversas de 203 estudantes distribuídas em sete atividades de programação de uma disciplina universitária de Inteligência Artificial. As mensagens dos estudantes foram mapeadas em atos de fala e modeladas como matrizes de transição de cadeias de Markov de primeira e segunda ordem no nível estudante-atividade. Após vetorização, padronização, redução de dimensionalidade com UMAP e clusterização com K-means, os grupos resultantes foram interpretados com apoio de métricas estruturais e modelos de processo extraídos com Heuristic*

Miner. A representação de segunda ordem produziu a solução mais coerente, com seis perfis de interação que variam de comportamentos rígidos e diretivos a padrões densamente exploratórios. Uma análise longitudinal baseada em distância de Hamming mascarada mostrou que muitos estudantes transitam entre perfis ao longo das atividades, em vez de permanecerem em um modo único e estável. Como principal contribuição, o estudo apresenta um pipeline explicável de learning analytics para dados conversacionais educacionais e mostra como esses rastros podem sustentar interpretações cautelosas, teoricamente informadas e relevantes para tutores por IA adaptativos.

Palavras-chave: learning analytics, aprendizagem autorregulada, tutoria por IA, atos de fala, mineração de processos.

1. Introdução

Sistemas de tutoria e agentes conversacionais mediados por Inteligência Artificial têm sido investigados como recursos para oferecer apoio personalizado em atividades de estudo, incluindo resolução de problemas, depuração de código e feedback sobre soluções [Labadze et al. 2023, Keuning et al. 2018]. Além da função pedagógica imediata, essas interações produzem registros conversacionais que podem subsidiar inferências sobre comportamentos e processos autorregulatórios dos estudantes [Winne 2017]. Em sistemas conversacionais, as mensagens de diálogo podem revelar modos de formular dúvidas, buscar ajuda, monitorar a própria compreensão e revisar estratégias ao longo da tarefa [Procter et al. 2018].

Apesar desse potencial, a maior parte dos estudos sobre comportamento de aprendizagem em ambientes digitais ainda se concentra em cliques, navegação, testes ou registros de submissão, e não em mensagens conversacionais com tutores baseados em grandes modelos de linguagem [Maldonado-Mahauad et al. 2018, Tang et al. 2023, Villalobos et al. 2024]. Essa lacuna é relevante porque os diálogos com tutores por IA tendem a se tornar cada vez mais frequentes em cenários educacionais, inclusive no ensino superior em Computação, e exigem abordagens analíticas que preservem interpretabilidade e conexão com a teoria educacional.

Neste trabalho, tratamos a aprendizagem autorregulada (*Self-Regulated Learning* – SRL) como lente interpretativa, e não como construto medido diretamente. Assim, atos de fala, reformulações, perguntas e padrões de transição são entendidos como *proxies* comportamentais compatíveis com processos como busca de ajuda, monitoramento e ajuste estratégico [Zimmerman and Schunk 2011, Roll et al. 2011, Azevedo et al. 2022, Poitras et al. 2023]. Essa postura reduz o risco de inferências excessivas sobre estados internos do estudante e favorece leituras mais prudentes de dados conversacionais.

O estudo utiliza o StudyChat, conjunto de dados público com diálogos autênticos entre estudantes e um tutor baseado em LLM (*Large Language Model*) em uma disciplina universitária de Inteligência Artificial [McNichols et al. 2026]. A contribuição é um pipeline explicável que combina atos de fala, cadeias de Markov, UMAP, K-means, métricas estruturais e mineração de processos para identificar perfis recorrentes e trajetórias longitudinais de interação.

O artigo responde a três perguntas de pesquisa:

1. Quais perfis recorrentes de interação emergem das sequências de atos de fala em diálogos com tutor por IA?
2. Em que medida esses perfis são interpretáveis quando analisados por métricas estruturais e modelos de processo?
3. Como os estudantes transitam entre perfis ao longo de atividades sucessivas?

2. Fundamentação e Trabalhos Relacionados

A literatura de SRL descreve a aprendizagem como processo no qual o estudante planeja, monitora, regula e avalia suas próprias ações em função de objetivos, dificuldades e feedbacks [Zimmerman and Schunk 2011]. Em ambientes digitais, entretanto, esses processos tendem a ser inferidos a partir de rastros de interação, e não observados como construtos internos diretamente acessíveis [Winne 2017]. Em vez disso, pesquisas em tutoria inteligente e *learning analytics* operam com manifestações comportamentais observáveis, como episódios de busca de ajuda, revisões, alternância de estratégias e reconfigurações ao longo da tarefa [Roll et al. 2011, Azevedo et al. 2022, Poitras et al. 2023].

Quando o objetivo é preservar a dinâmica temporal dessas manifestações, métodos sequenciais são especialmente adequados. Cadeias de Markov representam probabilidades de transição entre estados observáveis e permitem comparar perfis interacionais de modo compactado e interpretável. Modelos mais complexos, como modelos ocultos de Markov (*Hidden Markov Models* – HMMs), aprofundam a análise ao introduzir estados latentes, mas aumentam o custo interpretativo [Tang et al. 2023, Villalobos et al. 2024]. Neste artigo, priorizamos representações transparentes, pois o objetivo principal não é predição, mas interpretação educacional de perfis conversacionais.

A mineração de processos oferece uma segunda camada importante para essa interpretação. Em dados educacionais, ela tem sido usada para identificar caminhos recorrentes, gargalos, ciclos e configurações comportamentais associadas a estratégias de estudo [Maldonado-Mahauad et al. 2018]. De forma complementar, métodos de mineração sequencial também têm sido empregados para derivar padrões comportamentais relacionados a metacognição e uso de estratégias [Kinnebrew et al. 2013].

Trabalhos recentes reforçam essa direção metodológica em contextos mediados por IA generativa. Gao et al. analisam sequências de SRL em um sistema tutor baseado em IA generativa por meio de análise sequencial e clusterização [Gao et al. 2026]. De forma complementar, Wei et al. combinam análise de redes epistêmicas e mineração de padrões sequenciais para interpretar diálogos entre estudantes e sistemas de IA generativa em prática clínica, evidenciando a utilidade de abordagens temporais para compreender padrões de investigação em dados conversacionais [Wei et al. 2026].

3. Materiais e Métodos

3.1. Base de dados

O estudo utiliza o conjunto de dados público StudyChat [McNichols et al. 2026], construído a partir de diálogos autênticos entre estudantes e um tutor baseado em LLM em um curso universitário de Inteligência Artificial. O conjunto reúne 2.214 conversas, 16.851 mensagens e dados de 203 estudantes que consentiram em participar do estudo original, distribuídos em sete atividades de programação. O material passou por filtragem de informações pessoalmente identificáveis antes de sua disponibilização pública.

O contexto da base é relevante para interpretar os achados. Os estudantes podiam usar a ferramenta livremente durante as atividades, recorrendo ao tutor para esclarecer conceitos, discutir erros, pedir explicações sobre trechos de código ou revisar caminhos de solução. Essa liberdade de uso torna os diálogos mais próximos de situações autênticas de estudo mediado por IA, mas também amplia a heterogeneidade dos registros. Por essa razão, a unidade estudante-atividade foi adotada como forma de equilibrar granularidade analítica e comparabilidade entre casos.

A unidade de análise primária é o par estudante-atividade. Para cada estudante u e atividade t , as mensagens do estudante foram ordenadas temporalmente e convertidas em uma sequência discreta de atos de fala:

$$s^{(u,t)} = (s_1, s_2, \dots, s_T), \quad s_\tau \in \mathcal{A}. \quad (1)$$

Foram considerados apenas pares com ao menos duas mensagens do estudante. O foco exclusivo nas mensagens do estudante busca caracterizar seu comportamento interacional, e não a qualidade das respostas do tutor.

Embora o StudyChat disponibilize rótulos próprios de atos de fala em uma taxonomia de dois níveis, eles não foram utilizados neste artigo. Em seu lugar, as mensagens dos estudantes foram reclassificadas segundo as 41 categorias do *Switchboard Dialogue Act Corpus* (SwDA), adotadas como esquema fixo para a modelagem sequencial. Essa escolha emprega uma taxonomia consolidada em estudos de atos de fala e comparável a pesquisas recentes de classificação dialogal [Duran et al. 2023].

As classes do SwDA descrevem funções comunicativas recorrentes, como afirmações, perguntas, diretivas, confirmações e reformulações. Assim, o SwDA é a taxonomia usada para converter as mensagens em sequências discretas de atos de fala; 29 das 41 categorias ocorreram no corpus analisado.

Os artefatos de desenvolvimento preservados registram uma implementação supervisionada baseada no ajuste fino de `bert-base-uncased` [Devlin et al. 2019]. No experimento documentado, cada mensagem foi processada individualmente, com truncamento em 128 tokens, e um único rótulo SwDA foi selecionado pelo maior valor da camada *softmax*. Uma divisão estratificada de 90%/10% do próprio SwDA, com semente 42, reservou 19.974 enunciados para validação e registrou acurácia de 0,765 e macro-F1 de 0,437. Entretanto, o checkpoint e a saída integral de inferência que originaram os rótulos analisados não foram preservados, e não houve validação independente nas mensagens do StudyChat; essas métricas caracterizam apenas a avaliação interna no SwDA.

3.2. Representação sequencial

Cada mensagem do estudante foi, portanto, representada por um ato de fala do SwDA. Essa escolha também dialoga com estudos em domínios técnicos, nos quais atos de fala têm sido usados para caracterizar pedidos, explicações, comandos e respostas em conversas relacionadas a programação [Wood et al. 2018]. A análise trata esses rótulos como sinais observáveis de comportamento, não como medidas diretas de SRL.

Para cada sequência estudante-atividade, estimamos duas representações. Na pri-

meira ordem, a probabilidade de transição entre atos i e j é dada por:

$$P_{ij} = \frac{N(s_\tau = i \rightarrow s_{\tau+1} = j)}{\sum_{k \in \mathcal{A}} N(s_\tau = i \rightarrow s_{\tau+1} = k)}. \quad (2)$$

Na segunda ordem, a probabilidade do próximo ato considera os dois atos imediatamente anteriores, representados pelo par ordenado (a, b) :

$$P_{(ab) \rightarrow c} = \frac{N(s_{\tau-1} = a, s_\tau = b \rightarrow s_{\tau+1} = c)}{\sum_{k \in \mathcal{A}} N(s_{\tau-1} = a, s_\tau = b \rightarrow s_{\tau+1} = k)}. \quad (3)$$

As matrizes de transição foram vetorizadas por *flatten* e padronizadas com *z-score*. Em seguida, aplicamos *Uniform Manifold Approximation and Projection* (UMAP) para reduzir a dimensionalidade antes do agrupamento, pois essa técnica busca preservar estruturas locais e globais dos dados [Allaoui et al. 2020]. Sobre essa representação reduzida, usamos K-means para identificar grupos de sequências com dinâmica semelhante. O número de clusters foi selecionado por inspeção do método do cotovelo e confirmado por *silhouette*. A solução de primeira ordem indicou $k = 10$ com *silhouette* de 0,430, enquanto a solução de segunda ordem indicou $k = 6$ com *silhouette* de 0,464; a segunda foi mantida como solução principal por combinar maior coesão e melhor legibilidade interpretativa.

3.3. Interpretação dos clusters e trajetórias

A interpretação dos clusters combinou matrizes médias de transição, entropia, fator de ramificação, tamanho médio das sequências e modelos processuais gerados com Heuristic Miner, implementado no PM4Py 2.7.18 com os parâmetros padrão dessa versão [Maldonado-Mahauad et al. 2018]. Essa triangulação permitiu discutir quais perfis são mais compatíveis com busca de ajuda, monitoramento, execução estratégica ou ajuste metacognitivo, sem equiparar atos isolados a construtos psicológicos.

No plano longitudinal, a história de cada estudante foi representada pela sequência de clusters atribuídos às atividades:

$$c^{(u)} = (c_1, c_2, \dots, c_L), \quad c_\ell \in \{1, \dots, k\}. \quad (4)$$

Como nem todos os estudantes aparecem em todas as atividades, a similaridade entre trajetórias foi calculada por distância de Hamming mascarada, comparando apenas posições observadas em comum. Em seguida, trajetórias semelhantes foram agrupadas para identificar padrões recorrentes de estabilidade, alternância e mudança.

3.4. Operacionalização interpretativa de SRL e critérios de qualidade

Neste trabalho, SRL não foi inferida a partir de um ato de fala isolado, mas da combinação entre frequência, ordem temporal, estrutura processual e recorrência de comportamentos ao longo da atividade. Essa decisão metodológica evita tratar o traço observável como equivalente direto a um construto psicológico. Em vez disso, os clusters são descobertos primeiro pelos padrões transicionais; apenas depois são interpretados à luz da literatura de SRL, reduzindo circularidade analítica.

Também adotamos três critérios complementares para avaliar a qualidade da solução. O primeiro foi a coesão/separação estatística dos agrupamentos, observada por *silhouette*. O segundo foi a estabilidade interpretativa: procurou-se verificar se os perfis produziam descrições coerentes quando confrontados com matrizes médias, entropia, fator de ramificação e tamanho das sequências. O terceiro foi a legibilidade processual, isto é, a existência de caminhos dominantes compreensíveis nas redes extraídas pelo Heuristic Miner. Assim, a solução final não foi escolhida apenas por um indicador numérico, mas pelo equilíbrio entre qualidade estatística e utilidade explicativa para a interpretação educacional.

A Tabela 1 sintetiza as evidências observáveis empregadas na interpretação dos perfis e suas relações cautelosas com processos compatíveis com aprendizagem autorregulada.

Tabela 1. Evidências dialogais e leitura interpretativa adotada

Evidência observável	Leitura interpretativa	Processo compatível com SRL
Perguntas abertas e fechadas recorrentes	Busca ativa de entendimento e checagem de hipóteses	Busca de ajuda e monitoramento
Reformulações e sínteses parciais	Reorganização do raciocínio e explicitação do entendimento	Autoexplicação e autoavaliação parcial
Atos diretivos e orientados a passos	Foco em como prosseguir na tarefa	Planejamento e uso de estratégias
Ciclos entre perguntas, confirmações e reformulações	Ajuste iterativo do entendimento	Monitoramento e regulação adaptativa
Alta entropia e alto fator de ramificação	Exploração de alternativas e maior diversidade estratégica	Monitoramento distribuído
Baixa entropia e forte repetição de <i>Statement-non-opinion</i>	Execução rígida com baixa evidência de checagem	Execução com menor monitoramento

4. Resultados

4.1. Comparando as representações de primeira e segunda ordem

As duas representações testadas produziram agrupamentos com propriedades distintas. A modelagem de primeira ordem gerou dez clusters e permitiu observar grupos com repetição intensa de afirmações, alternância entre perguntas fechadas e respostas assertivas e dinâmicas mais difusas. Ainda assim, essa solução mostrou maior fragmentação do espaço analítico e menor capacidade de condensar contextos conversacionais mais ricos em um conjunto reduzido de perfis.

Na modelagem de segunda ordem, o condicionamento ao par de atos anteriores ampliou a sensibilidade do modelo ao histórico local da interação. Na prática, isso permitiu distinguir melhor situações em que os mesmos atos de fala apareciam, mas conectados por lógicas sequenciais distintas. Esse ganho explica por que a solução de seis clusters não apenas apresentou *silhouette* superior, mas também resultou em perfis mais claros para leitura pedagógica. Em outras palavras, a segunda ordem não apenas agrupou melhor, como também produziu unidades analíticas mais significativas para discutir monitoramento, busca de ajuda e execução estratégica.

A Figura 1 apresenta a projeção bidimensional dos vetores de segunda ordem e a distribuição dos seis agrupamentos identificados.

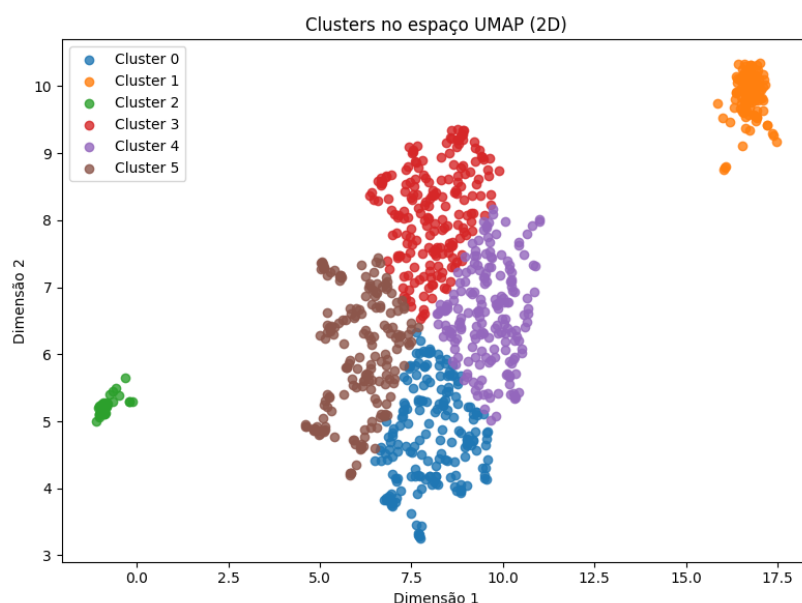


Figura 1. Projeção UMAP dos vetores derivados das matrizes de transição de segunda ordem, colorida pelos seis clusters da solução retida.

4.2. Perfis de interação intra-atividade

A solução final de segunda ordem produziu seis clusters interpretáveis. Em comparação com a representação de primeira ordem, a solução retida capturou melhor o contexto local do diálogo, reduzindo a fragmentação excessiva e preservando diferenças relevantes entre padrões conversacionais. Esse resultado reforça a importância de analisar o comportamento do estudante como uma organização temporal de ações, e não apenas como ocorrências isoladas, em consonância com estudos que utilizam padrões sequenciais e mineração de processos para interpretar estratégias de aprendizagem [Kinnebrew et al. 2013, Maldonado-Mahauad et al. 2018]. Nesse sentido, sequências curtas de atos de fala carregam informação adicional sobre estilo interacional e regulação da atividade. Na Tabela 2, comprimento médio corresponde à quantidade média de mensagens do estudante por interação estudante-atividade em cada cluster.

Para complementar a síntese agregada da Tabela 2, a Figura 2 mostra a distribuição das métricas estruturais recalculadas no nível estudante-atividade. Enquanto a tabela caracteriza cada perfil por valores agregados do cluster, os diagramas de caixa evidenciam

Tabela 2. Perfis finais de interação obtidos com matrizes de segunda ordem

Perfil	Mensagens	Entropia	Ramificação	Compr. médio	Leitura compatível com SRL
Exploratório / investigativo A	4005	1,585	6,375	11,779	Forte monitoramento e busca de ajuda
Localmente sistemático	339	1,527	3,526	2,306	Execução estruturada com checagens locais
Linear / diretivo	261	0,136	1,417	4,579	Predomínio de execução com baixa reflexão observável
Estratégia mista	3492	1,517	6,250	9,141	Alternância entre execução e monitoramento
Exploratório / investigativo B	4369	1,535	6,462	10,477	Busca de ajuda densa e regulação adaptativa
Semi-exploratório com caminhos preferenciais	3898	1,393	5,615	10,151	Execução estratégica com checagens recorrentes

a dispersão interna das unidades que compõem cada perfil. Essa distinção é importante porque entropia, fator de ramificação e comprimento médio na tabela descrevem o comportamento global do agrupamento, ao passo que a figura permite observar heterogeneidade, concentração e possíveis casos extremos.

O perfil linear/diretivo foi o agrupamento mais rígido, com entropia muito baixa, fator de ramificação reduzido e forte repetição de *Statement-non-opinion*. Essa classe do SwDA representa enunciados declarativos sem opinião explícita, isto é, afirmações factuais ou descritivas usadas para informar, relatar ou apresentar uma etapa da tarefa. Seu modelo processual mostrou fluxo fortemente concentrado nessa repetição, com poucas ramificações sustentadas. Em termos educacionais, esse padrão é compatível com situações nas quais o estudante permanece em modo de execução concentrada, com menos sinais observáveis de monitoramento metacognitivo.

Nos perfis exploratórios, ao contrário, as redes de processo exibiram maior densidade de arestas, ciclos frequentes e conexões entre afirmações, perguntas e outros movimentos dialogais. Nessas configurações, *Statement-non-opinion* aparece como nó central, mas não como ponto terminal do fluxo: ele recebe transições diversas e volta a distribuí-

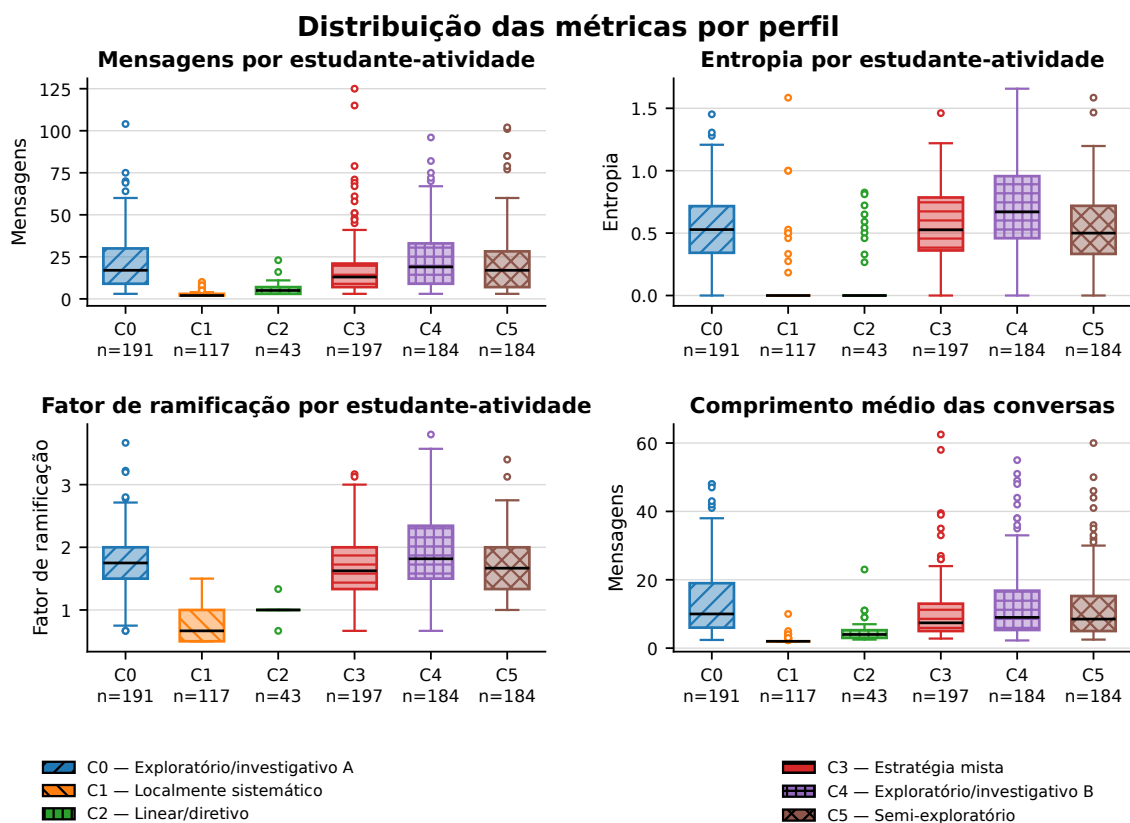


Figura 2. Distribuição das métricas estruturais no nível estudante-atividade. A Tabela 2 apresenta a síntese agregada por perfil, enquanto os diagramas de caixa mostram a dispersão interna das unidades atribuídas a cada cluster.

lo para novos questionamentos e reorganizações. Essa diferença ajuda a explicar por que contagens simples de atos de fala seriam insuficientes para distinguir sequências realmente exploratórias de sequências apenas repetitivas.

O perfil localmente sistemático ocupou posição intermediária. Seu comprimento médio curto e a ramificação moderada sugerem sequências mais contidas, mas ainda assim organizadas em torno de checagens localizadas e episódios pontuais de orientação. Já o perfil semi-exploratório preserva uma espinha dorsal de caminhos preferenciais, articulando perguntas e atos diretivos em torno de poucos eixos recorrentes, o que é compatível com execução estratégica acompanhada de monitoramento distribuído.

O perfil de estratégia mista, por sua vez, articulou momentos de condução mais diretiva com ciclos de pergunta, confirmação e reformulação. Em vez de um modo homogêneo de engajamento, ele parece capturar sequências nas quais o estudante alterna entre avançar na tarefa e reavaliar seu entendimento. Para *learning analytics*, essa distinção é especialmente relevante porque sinaliza que tutorias adaptativas não devem responder apenas ao tipo de ato atual, mas também ao padrão recente de transições.

4.3. Por que a mineração de processos foi necessária

Um resultado importante foi que as categorias não se diferenciaram apenas pela presença ou ausência de certos atos de fala, mas pelo modo como esses atos se conectavam em tra-

jetórias. O mesmo *Statement-non-opinion*, por exemplo, apareceu em vários clusters, mas com papéis processuais diferentes. No perfil linear/diretivo, ele concentrou a maior parte do fluxo em repetições, com menor diversidade de caminhos observada. Nos perfis exploratórios, ao contrário, funcionou como nó central que redistribui o fluxo para perguntas e novos ciclos de checagem.

Essa diferença mostra que, para dados conversacionais, a interpretação baseada apenas em frequências marginais seria insuficiente. A mineração de processos acrescentou poder explicativo ao tornar visíveis padrões de ramificação, retorno e densidade de caminhos. Para o campo de *learning analytics*, isso é especialmente relevante porque desloca a análise de uma lógica lexical ou distribucional para uma lógica processual, mais alinhada à ideia de aprendizagem como atividade temporalmente organizada.

O contraste apresentado na Figura 3 evidencia que atos de fala semelhantes podem ocupar papéis estruturais distintos nos dois perfis comparados.

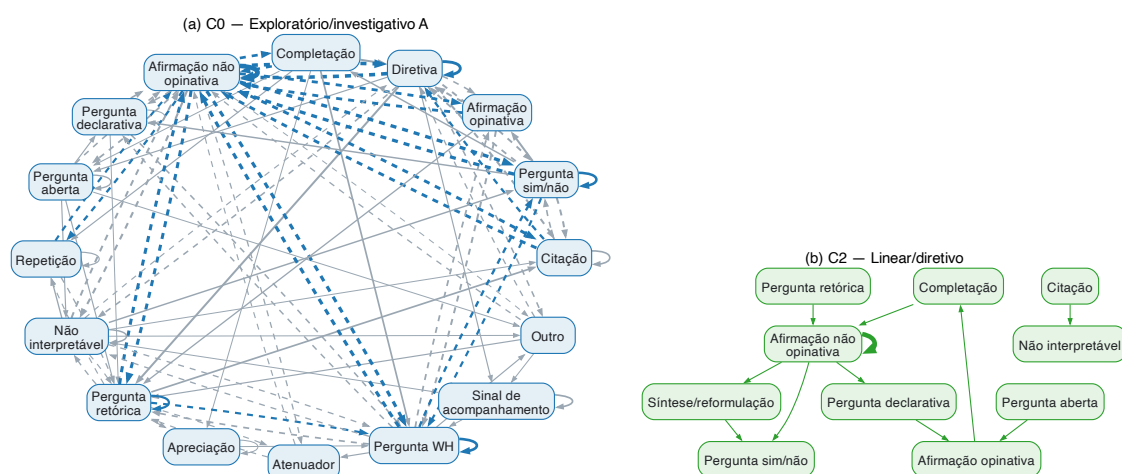


Figura 3. Heuristic Nets de C0 e C2, minerados com PM4Py e diagramados com Graphviz: (a) recorte de C0 com 16/24 atos e 96/117 arestas (99,2% das ocorrências); (b) C2 completo (10 nós, 11 arestas). Em C0, a frequência seleciona os nós (desempate pela frequência incidente), azul destaca as 25 arestas dominantes e cinza, as demais. A espessura indica frequência logarítmica e o tracejado, ciclos de comprimento dois; a mineração não foi alterada.

4.4. Trajetórias longitudinais

A análise longitudinal mostrou que os estudantes não permanecem presos a um único perfil de interação ao longo do curso. Muitos iniciam em configurações mais locais ou rígidas e depois migram para perfis mais amplos e exploratórios. Outros alternam sistematicamente entre perfis mistos e exploratórios, sugerindo reconfigurações em função da atividade, do tipo de dificuldade ou do momento da tarefa.

As trajetórias agrupadas por distância de Hamming mascarada formaram oito grupos com *silhouette* global de 0,2121. Esse valor é inferior ao observado nos clusters intra-atividade, o que era esperado diante da heterogeneidade longitudinal e da presença de sequências incompletas. Ainda assim, a análise mostrou três tendências recorrentes. A primeira foi a existência de estudantes relativamente estáveis em um perfil, sobretudo

nos agrupamentos exploratórios. A segunda foi a alternância regular entre perfis mistos e exploratórios. A terceira foi a ocorrência de episódios lineares/diretivos de curta duração, mais associados a demandas locais do que a um estilo persistente.

Esse resultado reforça uma interpretação importante para o desenho de sistemas educacionais: padrões interacionais devem ser tratados como configurações dinâmicas e dependentes do contexto, não como traços fixos do estudante. Essa leitura é consistente com estudos sobre aprendizagem autorregulada em ambientes tutoriais, que enfatizam a variação temporal de processos de monitoramento, regulação e uso de estratégias em função das demandas da tarefa [Azevedo et al. 2022, Poitras et al. 2023]. Em outras palavras, o mesmo estudante pode demandar tipos distintos de suporte em atividades diferentes.

A Figura 4 apresenta os fluxos entre os perfis ao longo das atividades e evidencia tanto permanências quanto mudanças de configuração.

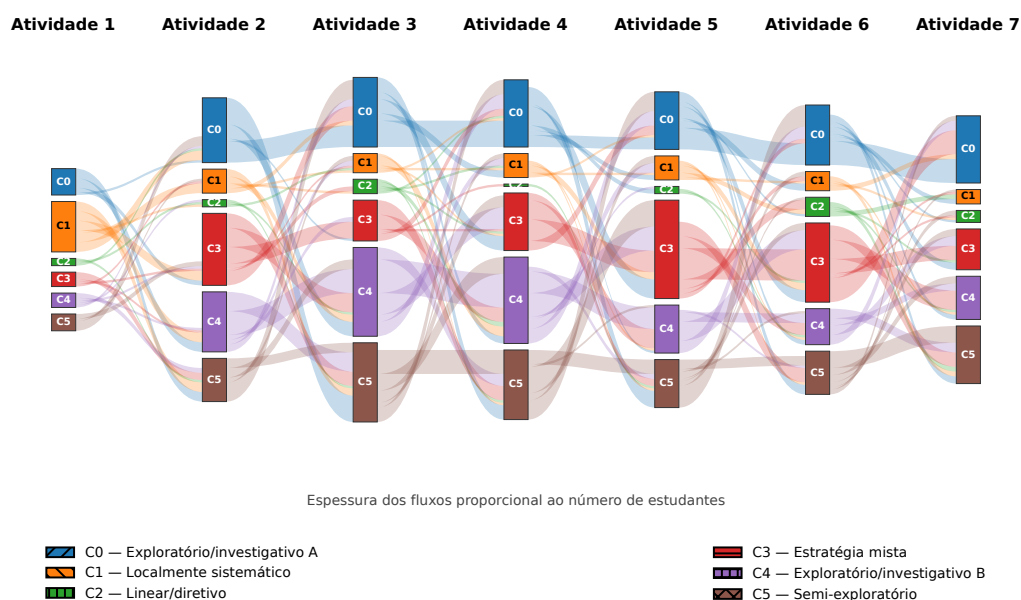


Figura 4. Transições longitudinais entre clusters de interação. Cada coluna representa uma atividade, e a espessura dos fluxos indica quantos estudantes aparecem em cada transição entre atividades consecutivas. As cores e a legenda associam os códigos dos clusters aos perfis interpretativos.

5. Discussão

Os achados contribuem para o campo de *learning analytics* em três direções: mostram que diálogos com tutores por IA podem ser analisados por métodos explicáveis; evidenciam que representações de segunda ordem capturam diferenças relevantes entre modos de interação que não aparecem em frequências marginais; e estendem a análise do nível intra-atividade para trajetórias longitudinais, em diálogo com estudos sobre manifestações de aprendizagem autorregulada ao longo de cursos [Villalobos et al. 2024].

Em termos de aplicação, os perfis sugerem que tutores por IA considerem, além do conteúdo da mensagem, a configuração interacional observada na atividade: configurações exploratórias e lineares/diretivas podem demandar apoios distintos, enquanto

transições entre atividades podem indicar momentos para intervenções metacognitivas ou autoavaliação. O pipeline produz saídas agregadas compreensíveis e trata dados conversacionais como processos temporalmente estruturados, oferecendo uma etapa interpretável antes de modelos preditivos mais complexos.

6. Ameaças à Validade

Em termos de validade de construto, atos de fala e padrões de transição são apenas manifestações observáveis do comportamento. Eles podem ser compatíveis com monitoramento, busca de ajuda ou execução estratégica, mas não autorizam inferência direta sobre estados psicológicos internos; por isso, o artigo discute compatibilidades interpretativas, sem afirmar que SRL foi medida.

Quanto à validade interna, os resultados dependem de escolhas analíticas específicas: taxonomia de atos de fala, critérios de filtragem, parametrização do UMAP, inicialização do K-means e representação apenas das mensagens do estudante. Outras combinações metodológicas podem produzir agrupamentos diferentes, embora a triangulação por métricas estruturais e modelos de processo reduza esse risco.

Os rótulos SwDA não foram submetidos a validação manual independente no StudyChat. Além do possível deslocamento de domínio entre as conversas telefônicas do SwDA e solicitações educacionais sobre programação, a ausência do checkpoint e da saída integral de inferência impede verificar se a execução avaliada corresponde exatamente aos rótulos processados. Erros sistemáticos de classificação podem, portanto, afetar as matrizes de transição e os perfis resultantes.

Em relação à validade externa, o StudyChat provém de um curso avançado, com estudantes familiarizados com programação e IA. Assim, os perfis não devem ser generalizados automaticamente para educação básica, formação docente, contextos de baixa infraestrutura ou domínios não técnicos.

7. Cuidados Éticos

O estudo realizou análise secundária de dados públicos disponibilizados após filtragem de informações pessoalmente identificáveis, sem contato direto com participantes, intervenção pedagógica ou coleta adicional de informações pessoais. A interpretação dos resultados foi mantida em nível agregado, sem tentativas de reidentificação e considerando os princípios gerais das Resoluções CNS 466/2012 e 510/2016.

8. Conclusão

Este artigo mostrou que Cadeias de Markov, UMAP, K-means e mineração de processos permitem organizar diálogos educacionais com tutores baseados em IA em perfis conversacionais interpretáveis. A representação de segunda ordem revelou seis perfis, de configurações lineares/diretivas a exploratórias, e a análise longitudinal identificou transições ao longo das atividades.

Como principal contribuição, o trabalho oferece um pipeline explicável para *learning analytics* em contextos educacionais com IA generativa e sustenta interpretações cautelosas, teoricamente informadas e sensíveis à temporalidade. Estudos futuros devem relacionar perfis e desempenho, incorporar o comportamento do tutor e replicar a análise em outros contextos.

Declaração sobre uso de Inteligência Artificial

Ferramentas de Inteligência Artificial generativa foram utilizadas como apoio à reorganização inicial do manuscrito e à revisão de redação, sempre sob supervisão humana integral. As decisões sobre recorte, método, interpretação dos resultados, redação final e responsabilidade pelo conteúdo permaneceram exclusivamente com as pessoas autoras.

Disponibilização de Artefatos

Os artefatos computacionais da análise estão disponíveis publicamente em <https://github.com/OtacilioN/artefatos-artigo-sbie-tutoria-ia>. O repositório reúne as sequências derivadas de atos de fala, matrizes e atribuições de clusters, métricas estruturais, artefatos de mineração de processos, trajetórias longitudinais, versões fixadas das dependências e scripts para validar os resultados e regenerar as figuras do artigo.

Referências

- Allaoui, M., Kherfi, M. L., and Cheriet, A. (2020). Considerably improving clustering algorithms using UMAP dimensionality reduction technique: A comparative study. In El Moataz, A., Mammass, D., Mansouri, A., and Nouboud, F., editors, *Image and Signal Processing: 9th International Conference, ICISP 2020, Marrakesh, Morocco, June 4–6, 2020, Proceedings*, volume 12119 of *Lecture Notes in Computer Science*, pages 317–325, Cham. Springer International Publishing.
- Azevedo, R., Bouchet, F., Duffy, M., Harley, J., Taub, M., Trevors, G., Cloude, E., Dever, D., Wiedbusch, M., Wortha, F., and Cerezo, R. (2022). Lessons learned and future directions of metatutor: Leveraging multichannel data to scaffold self-regulated learning with an intelligent tutoring system. *Frontiers in Psychology*, Volume 13 - 2022.
- Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In Burstein, J., Doran, C., and Solorio, T., editors, *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.
- Duran, N., Battle, S., and Smith, J. (2023). Sentence encoding for dialogue act classification. *Natural Language Engineering*, 29(3):794–823.
- Gao, J., Li, S., Zhang, J., Li, S., and Wang, T. (2026). Investigating self-regulated learning sequences within a generative ai-based intelligent tutoring system. In *Proceedings of the LAK26: 16th International Learning Analytics and Knowledge Conference*, LAK '26, pages 905–911, New York, NY, USA. Association for Computing Machinery.
- Keuning, H., Jeuring, J., and Heeren, B. (2018). A systematic literature review of automated feedback generation for programming exercises. *ACM Trans. Comput. Educ.*, 19(1).
- Kinnebrew, J. S., Loretz, K. M., and Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining*, 5(1).

- Labadze, L., Grigolia, M., and Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1):56.
- Maldonado-Mahauad, J., Pérez-Sanagustín, M., Kizilcec, R. F., Morales, N., and Muñoz-Gama, J. (2018). Mining theory-based patterns from big data: Identifying self-regulated learning strategies in massive open online courses. *Computers in Human Behavior*, 80:179–196.
- McNichols, H., Ikram, F., and Lan, A. (2026). The studychat dataset: Analyzing student dialogues with chatgpt in an artificial intelligence course. In *Proceedings of the LAK26: 16th International Learning Analytics and Knowledge Conference*, LAK '26, page 53–63, New York, NY, USA. Association for Computing Machinery.
- Poitras, E. G., Doleck, T., Huang, L., Dias, L., and Lajoie, S. P. (2023). Time-driven modeling of student self-regulated learning in network-based tutors. *Interactive Learning Environments*, 31(4):2490–2511.
- Procter, M., Lin, F., and Heller, B. (2018). Intelligent intervention by conversational agent through chatlog analysis. *Smart Learning Environments*, 5(1):30.
- Roll, I., Aleven, V., McLaren, B. M., and Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction*, 21(2):267–280.
- Tang, Y., Li, Z., Wang, G., and Hu, X. (2023). Modeling learning behaviors and predicting performance in an intelligent tutoring system: a two-layer hidden markov modeling approach. *Interactive Learning Environments*, 31(9):5495–5507.
- Villalobos, E., Pérez-Sanagustín, M., Azevedo, R., Sanza, C., and Broisin, J. (2024). Exploring manifestations of learners' self-regulated tactics and strategies across blended learning courses. *IEEE Transactions on Learning Technologies*, 17:1544–1557.
- Wei, J., Dang, D. K., Yang, K., Stokes, E., Mazeh, A., Lim, A., Dai, D. W., Moore, J., Fan, Y., Gasevic, D., Gasevic, D., and Chen, G. (2026). Uncovering students' inquiry patterns in genai-supported clinical practice: An integration of epistemic network analysis and sequential pattern mining. In *Proceedings of the LAK26: 16th International Learning Analytics and Knowledge Conference*, LAK '26, pages 261–271, New York, NY, USA. Association for Computing Machinery.
- Winne, P. H. (2017). Learning analytics for self-regulated learning. In Lang, C., Siemens, G., Wise, A., and Gašević, D., editors, *Handbook of Learning Analytics*, pages 241–249. Society for Learning Analytics Research.
- Wood, A., Rodeghero, P., Armaly, A., and McMillan, C. (2018). Detecting speech act types in developer question/answer conversations during bug repair. In *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2018, page 491–502, New York, NY, USA. Association for Computing Machinery.
- Zimmerman, B. J. and Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. In *Handbook of Self-Regulation of Learning and Performance*, pages 1–12. Routledge, New York, NY.