

AI Hype Labeling Guide (WSJ-Focused)

AI Hype Labeling Guide – WSJ-Focused, updated Version

This guide is designed for annotating articles from serious, measured outlets like The Wall Street Journal.

Focus on the **substance and implications** of the article rather than overt emotional tone.

You will assign two labels per article:

1. AI_RELEVANT: 1 (yes) or 0 (no) — is the article substantially about artificial intelligence?

2. If AI_RELEVANT = 1, assign a HYPE_LEVEL:

1 = Moderate hype, story about some AI-Application

2 = High hype, story about dramatic impact of AI

Step 1: Is the article about AI?

Label as AI_RELEVANT = 1 if:

- AI or machine learning is central to the article's topic (e.g., product launch, strategic shift, investment, regulation).
- The article covers companies, technologies, or policies explicitly involving AI.
- The focus is on AI's implications for business, markets, or society.

Step 2: Assign HYPE_LEVEL (only if AI_RELEVANT = 1)

1 = Moderate Hype / telling of a narrative

- Describes meaningful AI use with some optimism or concern.
- Shows growing importance of AI in a business or sector.
- Balanced discussion of opportunities and risks.
- Examples: AI improving workflows, enabling automation, is introduced / used by some company, or being tested at scale

2 = High Hype

- AI is framed as inevitable, transformative, society-wide or urgent.
- Language implies systemic change, large-scale disruption, or rapid adoption.
- Even if tone is neutral, the implications major societal or economic impact.
- A given stock price explodes due to AI-Boom
- Examples: "AI Overhauls Banking Sector", "Companies Race to Embed AI", "Investors Pour Billions into AI", "AI effects almost all sectors"

Remember: FEAR can also signal hype.

- If the article warns about massive job loss, policy panic, or existential threats from AI, this also counts as high hype.
- If AI should / might / will be regulated this is also signal for the relevance of AI

Quick Decision Guide:

- Ask: "Would this article make an investor more excited (or anxious) about the significance and urgency of AI?"
- If yes → likely Level 2.
- If cautious but clear significance or just an example of AI-usage in specific business → Level 1.

Tips:

- Use the dominant message of the article.
- If unsure, default to Level 1.