

## AI Hype Labeling Guide (WSJ-Focused)

AI Hype Labeling Guide – WSJ-Focused Version (3 Levels + AI Relevance Filter)

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:

0 = Low / None

1 = Moderate hype

2 = High hype

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.

Label as AI\_RELEVANT = 0 if:

- AI is mentioned only in passing or as a buzzword.
- The article is about a broader topic (e.g., markets, inflation) where AI is not the core subject.

## **Step 2: Assign HYPE\_LEVEL (only if AI\_RELEVANT = 1)**

0 = Low / No Hype

- Factual, technical, or skeptical tone.
- Reports slow adoption, limited impact, niche applications.
- Examples: AI research updates, infrastructure news, incremental improvements.

1 = Moderate Hype

- 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, 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 implication is major societal or economic impact.
- 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.

#### 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 → Level 1.

- If minimal or background context → Level 0.

#### Tips:

- Use the dominant message of the article.

- If unsure, default to Level 1.