



AI Engineer Case Study - Multi-Agent Systems

Your Mission: Build a multi-agent system using Pydantic AI that demonstrates your ability to design specialized AI agents that work together effectively.

Time Investment: 2-3 hours suggested (quality matters more than speed)

Deadline: 72 hours from receipt

Choose Your Focus Area

Option A: Customer Support Ticket Analyzer

Build a system with at least 2 specialized agents that analyze support tickets and determine appropriate routing.

Input Format:

```
{
  "ticket_id": "12345",
  "customer_tier": "premium", # free/premium/enterprise
  "subject": "API returning 500 errors intermittently",
  "message": "Hi, our production system has been failing...",
  "previous_tickets": 3,
  "monthly_revenue": 5000,
  "account_age_days": 450
}
```

Option B: Financial News Impact Analyzer

Build a system with at least 2 specialized agents that analyze financial news and assess market impact.

Input Format:

```
{
  "article_id": "news_001",
  "headline": "TechCorp announces record Q3 earnings...",
  "content": "TechCorp (NASDAQ: TECH) reported...",
  "published_at": "2024-10-15T14:30:00Z"
}
```

Core Requirements

1. Multi-Agent System

- Minimum 2 agents with distinct roles and perspectives
- Use Pydantic AI ([docs](#))
- Clear inter-agent communication
- **Think about:** What happens when agents have different opinions?

2. Evaluation Framework

- Test on all 5 provided test cases
- Implement at least 3 metrics
- Document performance
- **Consider:** How consistent are your agents?

3. Documentation

- System architecture - why multiple agents?
- Design decisions and rationale
- 2-3 prompt iterations showing improvement
- Examples of interesting test case behaviors

Deliverables

1. **Code:** Working implementation

2. **AI Chat Export:** Full conversation history showing your development process (as `ai_chat_history.txt`)
3. **Brief Write-up:** What didn't work and why (5-10 bullet points)
4. **Video (Optional but Recommended):** 3-5 minute explanation of your approach
 - Simple screen recording is fine
 - Candidates with videos get priority review

Evaluation Criteria

We're looking for:

- Clear thinking and problem-solving approach
- Ability to design truly specialized agents (not just one agent with different prompts)
- Understanding of multi-agent coordination - how do agents work together?
- Handling of edge cases - what happens in tricky situations?
- Quality of evaluation framework
- How you used AI tools effectively (shown via chat history)

Key Insight: The best solutions show agents that truly complement each other, not just divide work.

Test Data

Option A: Customer Support Tickets

Ticket SUP-001

```
{
  "ticket_id": "SUP-001",
  "customer_tier": "free",
  "subject": "This product is completely broken!!!",
  "message": "Nothing works! I can't even log in. This is the worst software I've ever used. I'm",
  "previous_tickets": 0,
  "monthly_revenue": 0,
  "account_age_days": 2
}
```

Ticket SUP-002

```
{
  "ticket_id": "SUP-002",
  "customer_tier": "enterprise",
  "subject": "Minor UI issue with dashboard",
  "message": "Hi team, just noticed the dashboard numbers are slightly misaligned on mobile view",
  "previous_tickets": 15,
  "monthly_revenue": 25000,
  "account_age_days": 730
}
```

Ticket SUP-003

```
{
  "ticket_id": "SUP-003",
  "customer_tier": "premium",
  "subject": "Feature Request: Bulk export",
  "message": "We need bulk export functionality for our quarterly reports. Currently exporting c",
  "previous_tickets": 5,
  "monthly_revenue": 5000,
  "account_age_days": 400
}
```

Ticket SUP-004

```
{
  "ticket_id": "SUP-004",
  "customer_tier": "premium",
  "subject": "API rate limits unclear",
  "message": "Getting rate limited but documentation says we should have 1000 requests/hour. We",
  "previous_tickets": 8,
  "monthly_revenue": 3000,
  "account_age_days": 180
}
```

Ticket SUP-005

```
{
  "ticket_id": "SUP-005",
  "customer_tier": "enterprise",
  "subject": "Urgent: Security vulnerability?",
  "message": "Our security team flagged that your API responses include internal server paths in error messages.",
  "previous_tickets": 20,
  "monthly_revenue": 50000,
  "account_age_days": 900
}
```

Option B: Financial News Articles

Article FIN-001

```
{
  "article_id": "FIN-001",
  "headline": "Tesla crushes Q3 expectations with record profits, but Musk warns of 'turbulent times' ahead",
  "content": "Tesla (NASDAQ: TSLA) reported stunning Q3 results with earnings of $1.05 per share, significantly exceeding analyst estimates. However, CEO Elon Musk cautioned investors about potential challenges in the coming quarters.",
  "published_at": "2024-10-22T16:00:00Z"
}
```

Article FIN-002

```
{
  "article_id": "FIN-002",
  "headline": "Small biotech CureGen soars on FDA approval, analysts remain skeptical",
  "content": "CureGen (NASDAQ: CURE), a small-cap biotech, received FDA approval for its novel cancer treatment, causing its stock price to surge over 200% in after-hours trading. Analysts, however, remain cautious about the company's long-term viability.",
  "published_at": "2024-11-01T14:30:00Z"
}
```

Article FIN-003

```
{
  "article_id": "FIN-003",
  "headline": "Amazon announces 'transformational' AI venture, but at massive cost",
  "content": "Amazon (NASDAQ: AMZN) unveiled Project Olympus, a $50 billion investment in AGI development, marking a pivotal shift in its AI strategy.",
  "published_at": "2024-09-15T09:00:00Z"
}
```

Article FIN-004

```
{
  "article_id": "FIN-004",
  "headline": "Regional bank FirstState posts record earnings amid industry turmoil",
  "content": "FirstState Bank (NYSE: FSB) reported record Q2 earnings of $3.20 per share, up 45% from last year, despite challenges in the regional banking sector.",
  "published_at": "2024-10-12T10:30:00Z"
}
```

Article FIN-005

```
{
  "article_id": "FIN-005",
  "headline": "China tech giant ByteDance reports stellar growth, regulatory clouds remain",
  "content": "ByteDance, TikTok's parent company, leaked financials show revenue grew 70% to $12 billion in Q3, though regulatory scrutiny in China and the US persists.",
  "published_at": "2024-11-21T18:45:00Z"
}
```

Submission Structure

```
your-name-case-study/
├── main.py
├── agents/
├── evaluation/
├── docs/
├── ai_chat_history.txt  # YOUR FULL AI CONVERSATION
└── README.md
```

Note: The test cases cover various scenarios - some straightforward, others more nuanced. Your system should handle the full spectrum.

Good luck!

We're looking for thoughtful engineers who value clarity and precision.