Kalshi

Algorithmic Trading Division

Cross Market Arbitrage

Capitalizing on pricing errors
between online prediction markets

Vienna, 21/01/2026

Polymarket

Meet the Team

Algorithmic Trading

Bence Pal

Associate

- Preprocessing
- Model Building
- Task Distribution

ERSTE Group

Amundi ASSET MANAGEMENT

- BSc. (CEU)
- MSc. (CEU) – 2nd Sem.

Christian Vorhauser

Senior Associate

- Polymarket Data Client
- Semantic Matching

ONB OESTERREICHISCHE NATIONALBANK EUROSYSTEM

s | p | v SPARKASSEN-PRÜFUNGSVERBAND

- BSc. BBE (WU)
- MSc. Qfin (WU) - 1st Sem.

Bartosz Rybarski

Fellow Analyst

- Real-time Data Processing
- Kalshi Data Client

amazon

STARKWARE

- BSc. CS (Uni. of Warsaw)
- M.Sc. SWE (TU) - 4th Sem

Leon Amon Höllhuber

Fellow Analyst

- Model Building
- Opportunity Detection

A1 Telekom Austria Group

Amundi Investment Solutions

- BSc. (UniWien)

Varvara Bondarenko

Fellow Analyst

- Preprocessing
- Pipeline Improvement

Deloitte.

- BSc. (UniWien)
- MSc. (TU) - 3rd Sem.

Agenda

Algorithmic Trading

1	**Project introduction**

2	**Methodology: technical infrastructure and toolset**

3	**Pipeline and semantic matching**

4	**Investment strategy: arbitrage execution**

5	**Conclusion and takeaways**

6	**Appendix**

Project overview

Cross-Market Arbitrage on Prediction Markets

Data	• Market data and real-time orderbook data from **Polymarket** and **Kalshi.**
Objective	• Develop a semi-automated pipeline to detect and capitalize on price discrepancies for **identical** contracts across platforms.
Approach	• **Automated Data Pipeline:** aggregation of live market data and standardization into a unified, high-speed structure. • **Semantic Market Matching:** linkage of equivalent events via AI-driven NLP to resolve naming inconsistencies across platforms.
Investment Strategy	• The system monitors **cross-market price spreads** to execute **simultaneous arbitrage trades**, capturing risk-free profits when discrepancies emerge between platforms
Tools	• **Polars:** execution of fast data normalization and high-volume cleaning. • **SentenceTransformers:** generation of dense vector embeddings for semantic analysis of market titles. • **Asyncio:** facilitation of efficient filtering and similarity-based market matching across platforms.

What is a Prediction Market?

Converting beliefs into tradable assets

The Idea and Motivation

Market beliefs

Prediction markets convert collective beliefs about future events into tradable prices.

Contract Value

Contracts usually represent binary events with a fixed payout, where the market price is driven by the event's assumed likelihood.

Arbitrage Gaps

Fragmentation across exchanges causes price divergence, enabling risk-free profit by securing opposing positions.

Example

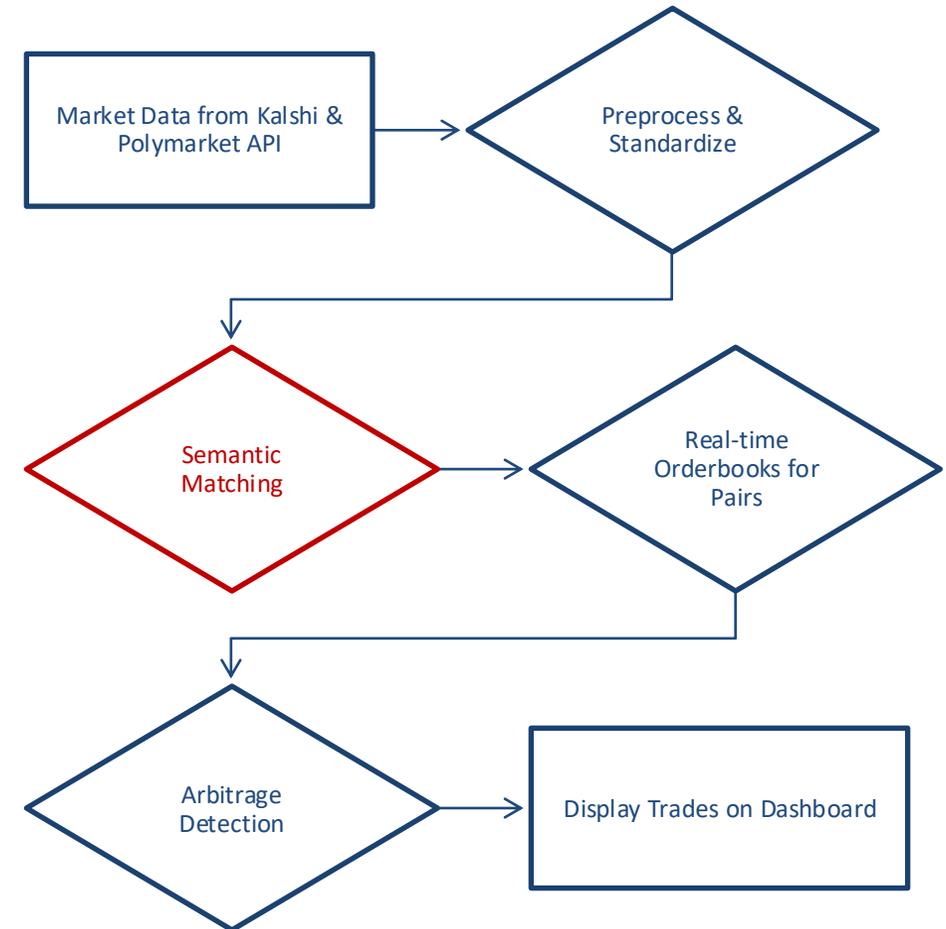**Arbitrage Scenario:** Polymarket sells "Yes" for 30¢, while Kalshi sells "No" for 66¢ for the same event.

Risk-Free Return: 96¢ total expenditure for a guaranteed $1.00 payout, locking in a 4.17% return.

Methodology

Simple Arbitrage Execution
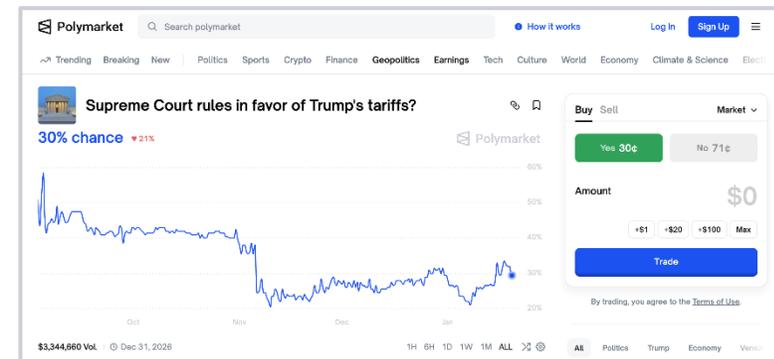
Data

- Polymarket: 1.6m markets, Kalshi: 2.2m markets
- After matching about 8500 pairs
- Most markets are very illiquid
- Symmetric Contracts: yes_ask = 1 - no_bid
- Vastly different format: Kalshi 1 ID, Polymarket 2 ID (for each outcome)

Pipeline Structure

```
Market Data from Kalshi &      ──►   Preprocess &
Polymarket API                        Standardize
                                          │
                                          ▼
Semantic Matching             ──►   Real-time Orderbooks for Pairs
        │
        ▼
Arbitrage Detection           ──►   Display Trades on Dashboard
```

Semantic Matching

From raw data to matched pairs

- **Normalize** raw market data
  - Standardized names, columns, magnitude, datatypes
- Extract unique market questions
- **Embed questions** using Sentence Transformer model
  - Small, effective model – not too computationally demanding
- Compute **cosine similarity** of all pairs
- Check with the **similarity threshold**
- Persist matches for arbitrage calculation
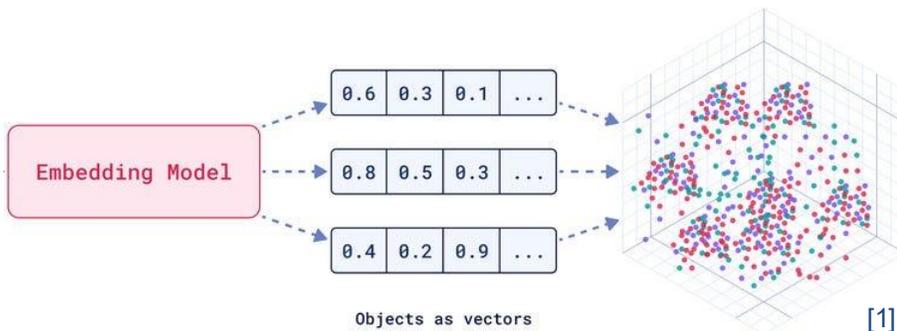
Matching Structure
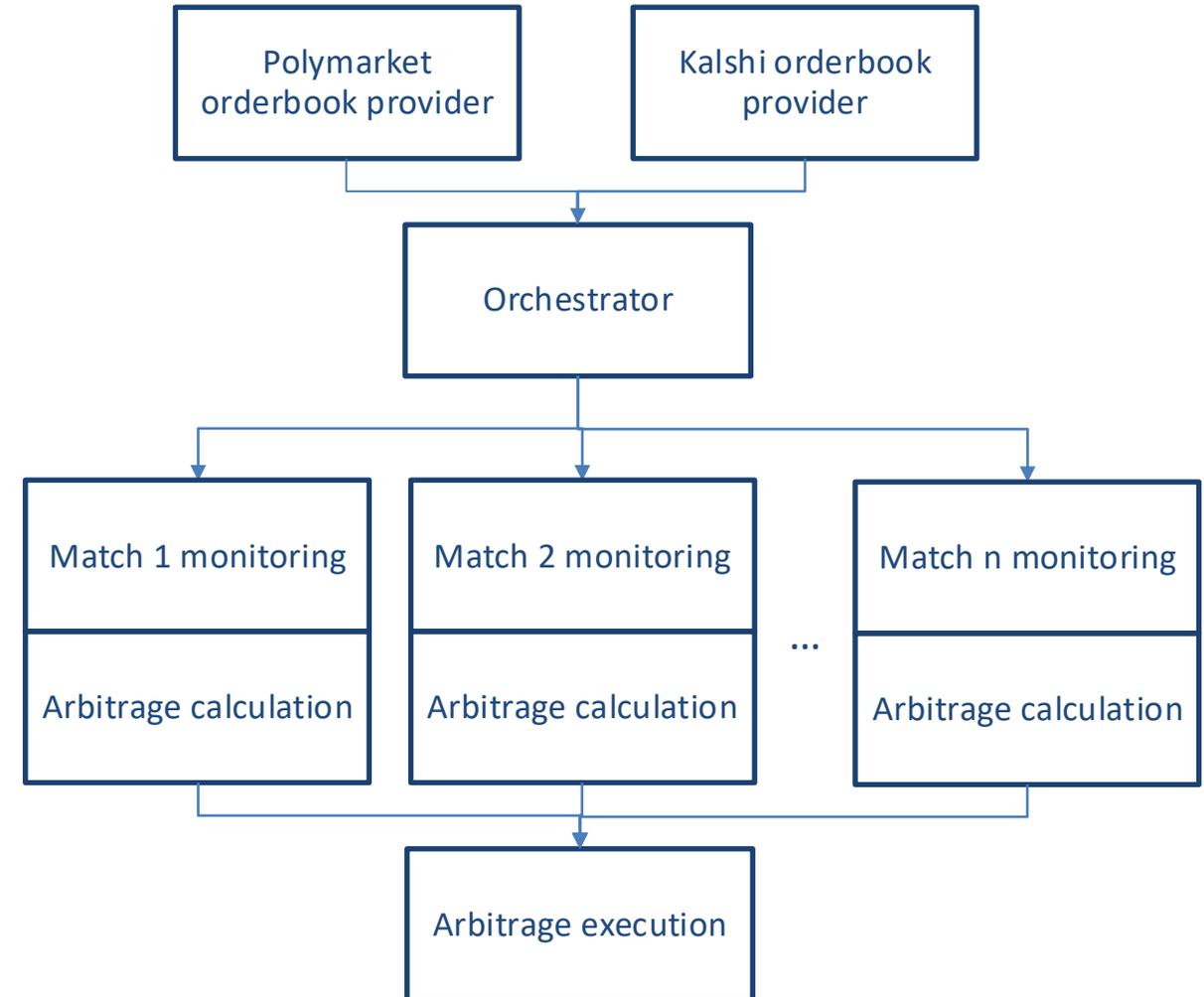
Realtime monitoring
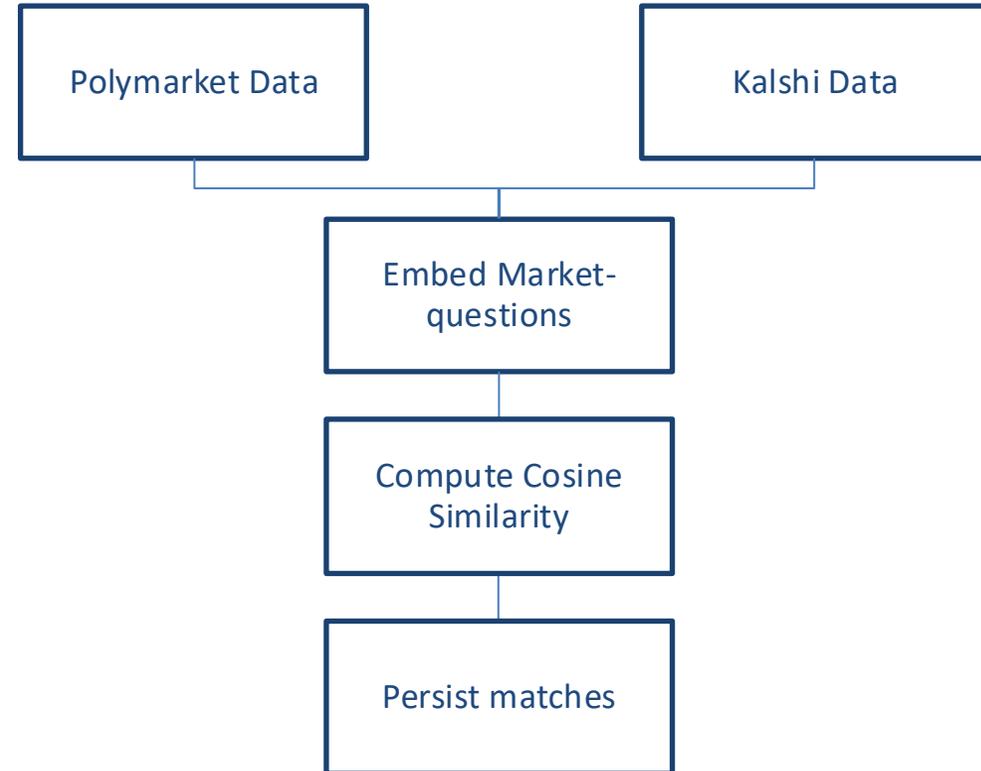
Fetching and evaluating orderbooks in real time.

Monitoring process

- Continuous querying of Polymarket's CLOB and Kalshi API for order books.
- Each matched pair gets a monitoring instance, which subscribes to orderbook providers, and continuously recalculates if there's an arbitrage opportunity.

Operational variants

- **High frequency monitoring**: Focus on a smaller set of matches, fetch orderbook with a very high frequency
- **Broad exploration**: Make rounds through all the matched pairs, but update orderbooks with lower frequency

```
┌──────────────────┐    ┌──────────────────┐
│    Polymarket     │    │  Kalshi orderbook │
│ orderbook provider│    │     provider      │
└──────────────────┘    └──────────────────┘
              │                 │
              └────────┬────────┘
                       ▼
              ┌──────────────────┐
              │   Orchestrator    │
              └──────────────────┘
                       │
        ┌──────────────┼──────────────┐
        ▼              ▼              ▼
┌───────────────┐ ┌───────────────┐ ┌───────────────┐
│Match 1        │ │Match 2        │ │Match n        │
│monitoring     │ │monitoring  ...│ │monitoring     │
├───────────────┤ ├───────────────┤ ├───────────────┤
│Arbitrage      │ │Arbitrage      │ │Arbitrage      │
│calculation    │ │calculation    │ │calculation    │
└───────────────┘ └───────────────┘ └───────────────┘
        └──────────────┬──────────────┘
                       ▼
              ┌──────────────────┐
              │Arbitrage execution│
              └──────────────────┘
```

Conclusion and Takeaways

High potential with broad room for improvement

Conclusion

- Existence of **time-adjusted post-fee** opportunities
- Arbitrage spreads widen for **mainstream events** with high volume and large amounts of retail investors
- **Huge opportunity:** even the simplest form of arbitrage can earn significant returns

Future Improvements

- Better contract matching
- Human feedback system
- Arbitrage against other platforms
- Using complex packaged contracts
- Statistical arbitrage

Filters

CSV path

arbitrage_opportunities.cs

Min expected profit

0.00   − +

Min quantity

0.00   − +

Max rows

200   − +

☐ Show fills column

Deploy ⋮

Arbitrage Opportunities 🔗

Opportunities	Total expected profit	Capital used	Latest
14	162.7232	18,083.2768	2026-01-21 01…

pair_name	timestamp	synthetic	total_quantity	avg_entry_cost	expected_profit
will montréal canadiens win the 202526 stanley cup finals	2026-01-21 01:36:17	KALSHI_YES__POLY_NO	5,605.11	0.9931	38.7803
will montréal canadiens win the 202526 stanley cup finals	2026-01-21 01:37:39	KALSHI_YES__POLY_NO	5,588.02	0.9931	38.6094
will tampa bay lightning win the 202526 stanley cup finals	2026-01-21 01:37:58	POLY_YES__KALSHI_NO	1,108.53	0.9865	14.9435
will tampa bay lightning win the 202526 stanley cup finals	2026-01-21 01:37:47	POLY_YES__KALSHI_NO	1,101.83	0.9865	14.8363
will edmonton oilers win the 202526 stanley cup finals	2026-01-21 01:36:23	KALSHI_YES__POLY_NO	1,389.58	0.99	13.8958
will tampa bay lightning win the 202526 stanley cup finals	2026-01-21 01:36:32	POLY_YES__KALSHI_NO	1,006.73	0.9868	13.3147
will toronto maple leafs win the 202526 stanley cup finals	2026-01-21 01:36:22	KALSHI_YES__POLY_NO	1,229.22	0.9901	12.2157
will south africa win the 2026 icc mens t20 world cup	2026-01-21 01:36:14	POLY_YES__KALSHI_NO	225	0.9764	5.3
will south africa win the 2026 icc mens t20 world cup	2026-01-21 01:38:10	POLY_YES__KALSHI_NO	175	0.9749	4.4
will the portugal win the 2026 mens world cup	2026-01-21 01:36:31	POLY_YES__KALSHI_NO	544	0.9922	4.2261

Thank you for your attention!

Any Questions?