

## Masked Candlestick Model: A Pre-Trained Model for Trading Prediction

**Authors :** Ling Qi, Matloob Khushi, Josiah Poon

**Abstract :** This paper introduces a pre-trained Masked Candlestick Model (MCM) for trading time-series data. The pre-trained model is based on three core designs. First, we convert trading price data at each data point as a set of normalized elements and produce embeddings of each element. Second, we generate a masked sequence of such embedded elements as inputs for self-supervised learning. Third, we use the encoder mechanism from the transformer to train the inputs. The masked model learns the contextual relations among the sequence of embedded elements, which can aid downstream classification tasks. To evaluate the performance of the pre-trained model, we fine-tune MCM for three different downstream classification tasks to predict future price trends. The fine-tuned models achieved better accuracy rates for all three tasks than the baseline models. To better analyze the effectiveness of MCM, we test the same architecture for three currency pairs, namely EUR/GBP, AUD/USD, and EUR/JPY. The experimentation results demonstrate MCM's effectiveness on all three currency pairs and indicate the MCM's capability for signal extraction from trading data.

**Keywords :** masked language model, transformer, time series prediction, trading prediction, embedding, transfer learning, self-supervised learning

**Conference Title :** ICDM 2023 : International Conference on Data Mining

**Conference Location :** Istanbul, Türkiye

**Conference Dates :** July 24-25, 2023