**Я пишу ВКР** по прогнозированию цен российских акций. Для этого необходимо проанализировать статьи. Здесь я как раз хочу рассказать про десяток статей, что я нашел на подобные темы, чем они отличались, что общего, на что опирались и какие результаты получили.

**Цели в статьях** были примерно следующие. Исследователи хотели прогнозировать цены акций с использованием различных методов таких как: регрессия, машинное обучение, эконометрика, анализ новостей и настроений. И старались разработать модели для повышения точности прогнозов и принятия инвестиционных решений.

Объекты были очень даже различны. Некоторые акцентировали своё внимание на какой-то отдельной бумаге, некоторые на нескольких и пытались сравнивать результаты, так же были различные фондовые индексы, например: S&P 500, NYSE, NASDAQ, AMEX или бразильский фондовый рынок. Ну и искали факторы, влияющие на их цены, такие как: финансовые новости, внешнеэкономические показатели, настроения инвесторов или изменение каких-то показателей кампаний и цен на сырье.

Гипотезы тоже отличались в зависимости от статей, например:

1. Методы машинного обучения (нейронные сети, случайный лес, LSTM) обеспечивают более точное прогнозирование цен акций по сравнению с традиционными методами (линейная регрессия, ARIMA).
2. Анализ новостей и настроений инвесторов улучшает точность прогнозирования.
3. Гибридные модели, сочетающие числовые данные и текстовый анализ, дают наилучшие результаты.
4. Современные алгоритмы и комбинированные модели превосходят традиционные методы в предсказании цен на фондовом рынке.
5. Также проверяются гипотезы об адекватности геометрического броуновского движения и линейных СДУ для описания движения цен акций.

**Основами для исследований** были традиционные экономико-математические методы (корреляционный, регрессионный анализ, ARIMA, GARCH), теории (эффективного рынка, случайного блуждания, поведенческая экономика) и современные методы машинного обучения. Учитывалось влияние новостей, общественного мнения мировых экономических показателей, мультипликаторов и волатильности. Использовались стохастические модели, модели геометрического броуновского движения, модели Марковского переключения и методы разложения сигналов.

**Данные исследования** очень важны ведь есть необходимостью повышения точности прогнозирования цен акций для минимизации рисков и увеличения доходности инвестиций в условиях волатильности рынков, нестабильной экономики и глобализации. Конечно, существующие методы прогнозирования несовершенны, особенно в условиях увеличения объема данных и глобализации. Исследования направлены на восполнение пробелов путем интеграции новых подходов: математических и статистических методов, машинного обучения, нейронных сетей, гибридных моделей, анализа новостей и учета поведенческих факторов. Особое внимание уделяется применению этих методологий на конкретных рынках, включая развивающиеся, где наблюдается дефицит исследований. Результаты важны для инвесторов и участников рынка, стремящихся принимать обоснованные решения в условиях неопределенности и меняющейся экономической ситуации, а также растущей популярности биржевых торгов. Развитие вычислительных мощностей и доступности данных делает эти методы стандартом прогнозирования на фондовых рынках в будущем. Новые модели прогнозирования будут полезны в долгосрочной перспективе.

**Авторы статей ставят задачи**: прогнозирование цен акций с использованием различных методов (линейная регрессия, ANN, RF, ARIMA, LSTM и др.), анализ факторов, влияющих на точность прогнозирования, сравнение эффективности методов, разработка и тестирование моделей, в том числе гибридных и на основе машинного обучения, с использованием временных рядов, новостных данных и других факторов. Задачи включают также оценку целесообразности использования текстовой информации, анализ волатильности, корреляционный анализ, проверку мультиколлинеарности, оптимизацию моделей, выявление преимуществ и недостатков подходов, разработку рекомендаций и алгоритмов, а также анализ влияния поведенческих особенностей инвесторов и рыночных паттернов.

**Данные брались из многих** источников, Московская Биржа, Quandl, Finam.ru, Tushare и другие. Использовались дневные и минутные значения, высокие, низкие, открытые и закрытые цены, объемы торгов и смещенные значения. Также использовались экономические индикаторы (цены на нефть, курс доллара, ВВП Китая), финансовые новости из LexisNexis, Reuters, Bloomberg, Kaggle и других источников, текстовые данные из социальных сетей (форумы, Twitter) и финансовые отчеты компаний. Анализировались данные по различным компаниям (Газпром, Apple, Сбербанк и др.), индексам (S&P 500, IHSG), а также отраслевые индексы. Данные собираются через API, веб-скраппинг и из аналитических отчетов.

**Но в статьях так же** говорилось о недостатках:

Например, недостаточное количество компаний, короткий временной период, узкий географический охват, ограниченный набор переменных или метрик, зависимость от исторических данных. Где-то не учитываются макроэкономические факторы, новостной фон, транзакционные издержки, рыночное воздействие, изменения в рыночной ситуации. В некоторых статьях есть зависимость от исторических данных, сложность настройки и интерпретации моделей, ограниченность в типах используемых моделей и подходов, отсутствие гибридных решений. Нету данных о применимости в реальных условиях, ограничение для краткосрочного прогнозирования, отсутствие оценки влияния внешних факторов на точность, специфичность результатов. Ну и не рассматриваются долгосрочные тренды, медвежьи рынки, реальные торговые стратегии, ограничения исследования и возможные ошибки в прогнозах.

**В статьях для оценки моделей** прогнозирования цен акций применялись метрики RMSE, MSE, MAPE, R², коэффициенты корреляции для измерения точности и качества прогнозов. Для классификации использовались Precision, Recall, F1-score, Accuracy, MCC, TNR. Финансовые показатели (доходность), коэффициенты дрейфа и волатильности, а также информационные критерии AIC и SC также применялись. Подчёркивалась важность метрик для количественной оценки расхождений и эффективности моделей. MSE, MAPE, MDA оценивали общую ошибку и правильность определения направления изменения. AIC и SC использовались для проверки значимости эффектов в моделях.

**В статьях для** прогнозирования цен акций использовались разнообразные модели: линейная регрессия, искусственные нейронные сети (ANN), LSTM, RNN, GRU, MLP, случайный лес, SVM, деревья решений, ARIMA, GARCH, SVR, MKL, геометрическое броуновское движение, одно- и многослойные нейронные сети, FB Prophet и другие.

**LSTM часто демонстрировала** наилучшую точность. Успешно применялись гибридные модели, а также комбинации различных моделей и данных из разных источников (например, новостей).

* Отдельные модели показывали точность до 96%.
* ANN превосходит RF, LSTM лучше SVM, KNN, PCA и обычной нейросети.
* LSTM и SVR эффективны для краткосрочного прогнозирования.
* Новости на португальском позволяют предсказывать падения цен.
* Использование всех категорий новостей повышает точность до 82.40%.
* ANN точнее HMM и SVR, DWT-ANN и EMD-SVR улучшают прогнозы волатильных индексов.
* Геометрическое броуновское движение точно для краткосрочных прогнозов.
* LSTM превосходит ARIMA.
* Выявлена корреляция цен 'Лукойл' и нефти Brent.
* LSTM демонстрировал низкий RMSE с новостной информацией.
* Однослойная нейросеть обеспечивает высокую точность прогноза.
* Использование метода расширения признаков и корректировка LSTM улучшили показатели моделей.
* Модели LSTM и GRU достигают наименьшей средней квадратичной ошибки с 5 скрытыми слоями.
* Объединенная модель с анализом новостей достигла минимальной ошибки.

**Практическая значимость исследования** заключается в улучшении точности прогнозирования фондового рынка и цен акций. Это может быть полезно для инвесторов, трейдеров, финансовых аналитиков и регуляторов при принятии обоснованных инвестиционных решений, минимизации рисков и предотвращении финансовых потерь. Предложенные модели и методики могут использоваться для:

• Поддержки решений на биржевых рынках.

• Предоставления информированных рекомендаций на основе анализа новостей.

• Оценки инвестиционной привлекательности ценных бумаг.

• Создания эффективных стратегий торговли.

• Улучшения финансовых предсказательных систем.

• Прогнозирования финансово-экономических показателей.

• Определение стратегии торговли.

• Улучшения качества инвестиционных решений.

Таким образом, исследования в этом направлении могут помочь людям быть спокойнее за свои накопления и будущее. А компаниям быть увереннее в бизнесе и своей живучести на рынке.

Спасибо за внимание.