

УДК 004.89

## Прогнозирование XRD спектров тонких $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$ пленок методами машинного обучения

Сапего Евгений Николаевич

eugenysapego@yandex.ru

SPIN-код: 980-246

Закасовский Игорь Николаевич

igorzakasovskij1@gmail.com

СПбГЭТУ «ЛЭТИ», Санкт-Петербург, Россия

**Аннотация.** Показана возможность прогнозирования дифрактограмм сегнетоэлектрических тонких пленок  $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$  методами машинного обучения в зависимости от технологических параметров магнетронного ВЧ-распыления, состава мишени и типа подложки. Проведен сравнительный анализ моделей, выявлена наилучшая по точности модель Random Forest и рассмотрены способы повышения достоверности прогнозов путем нормализации и обработки данных.

**Ключевые слова:** машинное обучение, сегнетоэлектрики, тонкие пленки, магнетронное ВЧ-распыление, рентгеноструктурный анализ

## Machine - Learning Prediction of X - ray Diffraction Patterns of $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$ Thin Films

Sapego Evgeny Nikolaevich

eugenysapego@yandex.ru

SPIN-code: 980-246

Zakasovsky Igor Nikolaevich

igorzakasovskij1@gmail.com

Saint Petersburg Electrotechnical University, Saint-Peterburg, Russia

**Abstract.** The study demonstrates the possibility of predicting X-ray diffraction (XRD) patterns of ferroelectric  $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$  thin films using machine-learning methods depending on the technological parameters of RF magnetron sputtering, target composition, and substrate type. A comparative analysis of models was carried out, showing that the Random Forest algorithm provides the highest prediction accuracy, while data normalization and preprocessing significantly improve model performance.

**Keywords:** machine learning, ferroelectrics, thin films, magnetron sputtering, X-ray diffraction analysis

**Введение.** Тонкие сегнетоэлектрические (СЭ) пленки широко применяются при создании компонентов СВЧ-электроники [1–4]. Электрофизические характеристики таких структур — управляемость емкости и величина диэлектрических потерь — определяются их кристаллическим строением, зависящим от технологических параметров осаждения. Следовательно, актуальной задачей является выбор оптимальных режимов синтеза, обеспечивающих требуемые структурные свойства материала.

Традиционно оптимизация параметров осуществляется экспериментальным путем [5, 6], что требует значительных временных и материальных затрат при выполнении серий многофакторных экспериментов. Современным направлением в материаловедении, позволяющим ускорить разработку материалов с заданными свойствами, является применение методов машинного обучения (МО). Эффективность данного подхода показана в работах, посвященных прогнозу свойств материалов [7–10], оптимизации технологий синтеза [11–13], поиску новых соединений [14–15].

Целью настоящего исследования является применение моделей МО для предсказания рентгеновских дифрактограмм тонких пленок титаната-станната бария ( $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$ ), полученных методом ВЧ-магнетронного распыления. Такой подход позволяет установить связь между технологическими параметрами процесса и структурными характеристиками сформированных пленок.

**Материалы и методы.** Тонкие сегнетоэлектрические пленки состава  $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$  (BSnT) были получены методом ВЧ-магнетронного распыления при различных технологических режимах. Структурные свойства исследовались методом рентгеновской дифракции на дифрактометре ДРОН-6 с использованием излучения  $\text{CuK}\alpha_1$  ( $\lambda = 1,5406 \text{ \AA}$ ). В табл. 1 представлены диапазоны значений основных технологических параметров, состава твердых растворов и типов подложек.

Таблица 1

Значения технологических параметров

| Признак                | Диапазон значений                                                       |
|------------------------|-------------------------------------------------------------------------|
| Давление рабочего газа | 2...10 Па                                                               |
| Время осаждения        | 80...360 мин                                                            |
| Состав рабочего газа   | $\text{Ar}/\text{O}_2 = 0...100 \%$                                     |
| Температура подложки   | 650...900 °C                                                            |
| Состав мишени          | $\text{BaSn}_x\text{Ti}_{(1-x)}\text{O}_3$ ( $x = 0,2; 0,3; 0,4; 0,5$ ) |
| Тип подложки           | сапфир, поликор, SiC                                                    |

В качестве целевой переменной рассматривалась интенсивность рентгеновского излучения в зависимости от угла дифракции  $2\Theta$  в диапазоне 20–60°, соответствующем наиболее информативной области дифракционных пиков. Общий объем данных составил 148 образцов, разделенных на обучающую (80 %) и валидационную (20 %) выборки.

Для построения прогностической модели использовались алгоритмы Linear Regression (LR), Regression Tree (RT), Random Forest (RF) и LightGBM (LGBM). Для повышения точности выполнена min-max нормализация целевой переменной и стандартизация признаков. Качество моделей оценивалось по значениям метрик RMSE, MAE,  $R^2$  и коэффициента Пирсона ( $\rho$ ).

На первом этапе был выполнен поиск оптимальной модели по метрикам ошибок на валидационной выборке. Отобранная модель применялась для

прогнозирования дифрактограмм тонких СЭ-пленок  $\text{BaSn}_{0,2}\text{Ti}_{0,8}\text{O}_3$ , осажденных на подложках из поликора, сапфира и SiC при различных технологических параметрах. На основе предсказаний определены оптимальные режимы осаждения, после чего выполнен экспериментальный синтез и последующий рентгеноструктурный анализ для сопоставления полученных и прогнозируемых дифрактограмм.

**Результаты.** Численные значения метрик ошибок на валидационной выборке представлены в табл. 2. Из приведенных данных видно, что модели, основанные на деревьях решений, демонстрируют существенно более высокую точность по сравнению с линейной регрессией. Наилучшие показатели достигнуты для алгоритма Random Forest, обеспечивающего значения  $R^2 = 0,945$  и  $\rho = 0,968$ .

Таблица 2

Числовые значения функций ошибок на валидационной выборке

| Модель                 | Функция ошибок |        |         |         |
|------------------------|----------------|--------|---------|---------|
|                        | $R^2$          | $\rho$ | RMSE    | MAE     |
| Random Forest (RF)     | 0,945          | 0,968  | 222,154 | 38,993  |
| Regression Tree (RT)   | 0,920          | 0,957  | 268,651 | 45,149  |
| LightGBM (LGBM)        | 0,770          | 0,882  | 460,037 | 124,335 |
| Linear Regression (LR) | 0,012          | 0,172  | 952,777 | 167,846 |

Экспериментальная проверка точности модели RF проведена для пленок  $\text{BaSn}_{0,2}\text{Ti}_{0,8}\text{O}_3$ , осажденных на различных типах подложек. Сравнение фактических и прогнозируемых дифрактограмм приведено на рис. 1–3.

Метрики ошибок тестовой проверки представлены в табл. 3. Высокие значения коэффициентов детерминации ( $R^2 > 0,9$ ) и корреляции ( $\rho > 0,95$ ) подтверждают надежность прогноза модели для различных типов подложек.

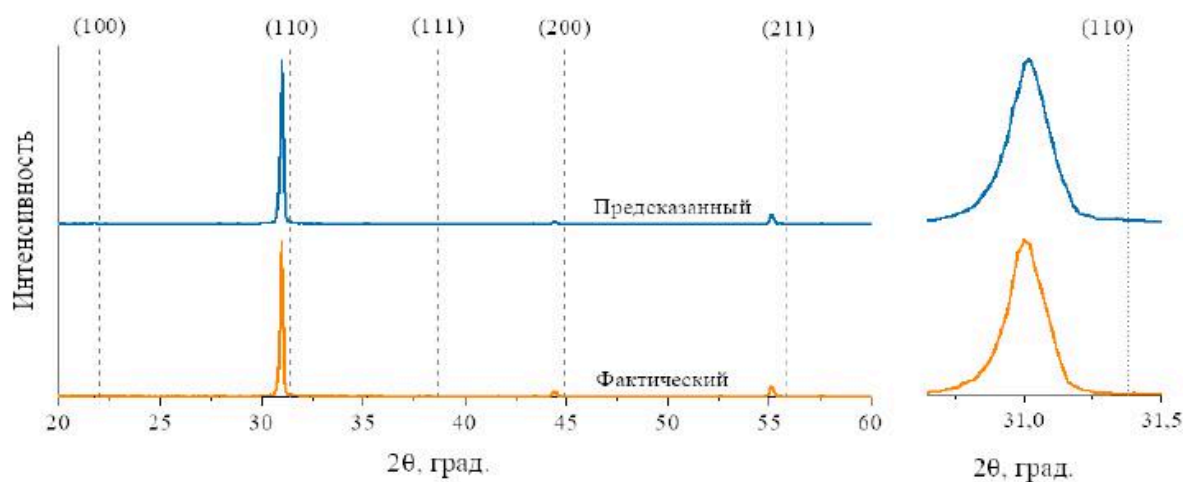


Рис. 1. Сравнение фактической и предсказанной дифрактограмм тонкой пленки, твердого раствора  $\text{BaSn}_{0,2}\text{Ti}_{0,8}\text{O}_3$  на подложке из поликора

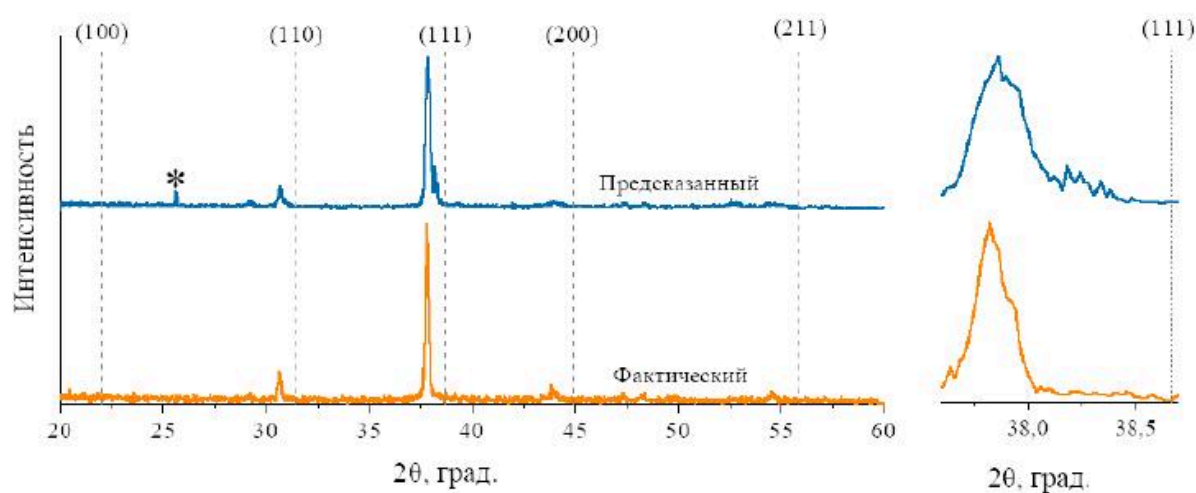


Рис. 2. Сравнение фактической и предсказанной дифрактограмм тонкой пленки, твердого раствора  $\text{BaSn}_{0.2}\text{Ti}_{0.8}\text{O}_3$  на подложке из сапфира

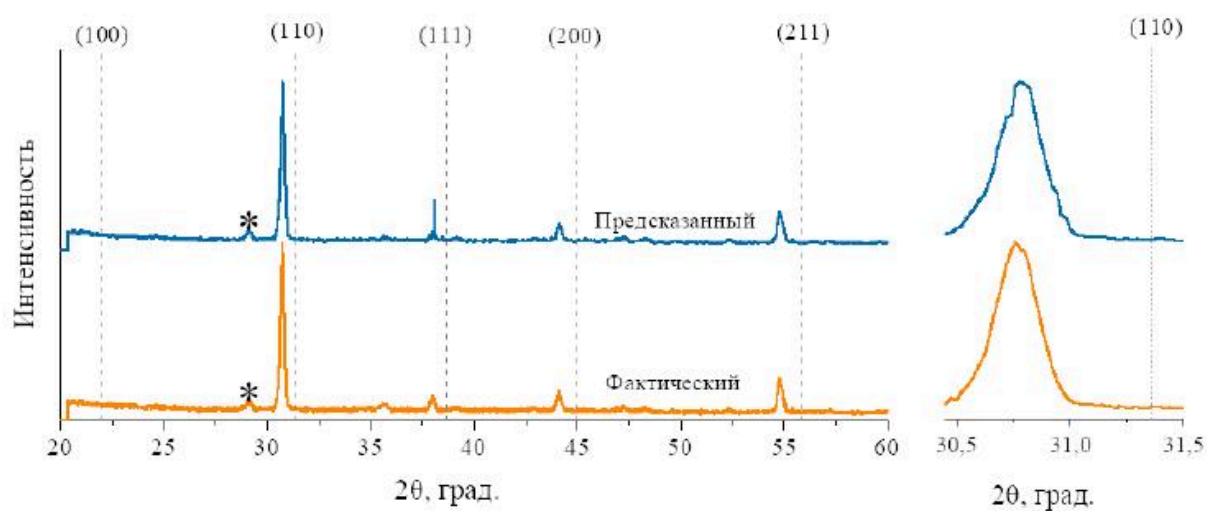


Рис. 3. Сравнение фактической и предсказанной дифрактограмм тонкой пленки, твердого раствора  $\text{BaSn}_{0.2}\text{Ti}_{0.8}\text{O}_3$  на подложке из SiC

Таблица 3.

Числовые значения функций ошибок для тестовой проверки

| Подложка | Функция ошибки |        |        |        |
|----------|----------------|--------|--------|--------|
|          | $R^2$          | $\rho$ | RMSE   | MAE    |
| Поликор  | 0,980          | 0,994  | 66,910 | 16,370 |
| Сапфир   | 0,905          | 0,952  | 21,870 | 11,234 |
| SiC      | 0,985          | 0,993  | 51,518 | 24,095 |

**Обсуждение результатов.** Сравнительный анализ показал, что модели, основанные на алгоритмах деревьев решений (Random Forest, Regression Tree, LightGBM), обеспечивают наилучшую точность предсказания по всем метрикам. Наиболее высокие значения коэффициентов детерминации ( $R^2 = 0,945$ )

и корреляции ( $\rho = 0,968$ ) получены для Random Forest, что связано со способностью ансамблевых методов эффективно аппроксимировать нелинейные многофакторные зависимости и устойчиво работать с экспериментальными данными ограниченного объема. В отличие от них, линейная регрессия (LR) не отражает сложную структуру зависимости интенсивности рентгеновского сигнала от технологических параметров, что приводит к значительному снижению точности.

Экспериментальная проверка подтвердила высокую надежность прогноза дифрактограмм на различных типах подложек. Наиболее полное совпадение между фактическими и рассчитанными данными наблюдается для образцов, осажденных на поликоре и SiC, что указывает на корректное моделирование влияния текстуры и дефектности кристаллической решетки. Несколько меньшая точность прогноза для сапфировой подложки может быть обусловлена неоднородностью обучающей выборки и особенностями роста пленки на монокристаллическом оксиде алюминия.

Таким образом, применение алгоритмов Random Forest и близких к нему методов позволяет достоверно прогнозировать рентгеноструктурные характеристики сегнетоэлектрических пленок  $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$  и использовать модельные данные для предварительной оптимизации технологических режимов осаждения.

**Заключение.** Проведенное исследование показало, что применение моделей машинного обучения, в частности алгоритма Random Forest, позволяет достоверно прогнозировать рентгеновские дифрактограммы тонких пленок  $\text{BaSn}_x\text{Ti}_{1-x}\text{O}_3$  в зависимости от параметров ВЧ-магнетронного распыления. Полученные результаты демонстрируют возможность целенаправленной оптимизации технологических режимов осаждения и управления структурными свойствами сегнетоэлектрических покрытий, что открывает перспективы для разработки функциональных СВЧ-компонентов с заданными характеристиками.

*Работа выполнена при финансовой поддержке Министерства науки  
и высшего образования Российской Федерации в рамках гранта  
№ 075-01438-22-07 — FSEE -2025-0010.*

## Список источников

- [1] Алтынников А.Г. и др. Характеризация свойств пленок титаната бария-стронция и управляемых элементов на их основе в диапазоне частот 1–60 GHz. *Письма в Журнал технической физики*, 2019, т. 45, № 11, с. 12–15.
- [2] Miranda F. A. et al. Design and development of ferroelectric tunable microwave components for Ku and K-band satellite communication systems. *IEEE Transactions on Microwave Theory and Techniques*, 2000, vol. 48, no. 7, pp. 1181–1189.
- [3] Мироненко И.Г. и др. Щелевая линия на основе нанокompозитных сегнетоэлектрических пленок. *Радиотехника*, 2012, № 7, с. 117–122.
- [4] Вендик О.Г., Козырев А.Б. Применение сегнетоэлектриков в антеннах с электронным сканированием. *Известия высших учебных заведений России. Радиоэлектроника*, 2005, № 1, с. 20–23.

- [5] Коваленко Д.А., Петров В.В., Клиндухов В.Г. Исследование влияния технологических параметров формирования тонких пленок цирконата-титаната свинца на их структурные и электрофизические свойства. *Известия Южного федерального университета. Технические науки*, 2014, № 9 (158), с. 124–132.
- [6] Букарев И.М., Жданов А.В. Исследование влияния технологических параметров процесса магнетронного осаждения на износостойкость покрытия CrN/AlN. *Современные проблемы науки и образования*, 2012, № 2, с. 277–277.
- [7] Wang L., Mu Z., Guo H. Application of support vector machine in the prediction of mechanical property of steel materials. *Journal of University of Science and Technology Beijing, Mineral, Metallurgy, Material*, 2006, vol. 13, no. 6, pp. 512–515.
- [8] Kwak S. et al. Machine learning prediction of the mechanical properties of  $\gamma$ -TiAl alloys produced using random forest regression model. *Journal of Materials Research and Technology*, 2022, vol. 18, pp. 520–530.
- [9] Li X. et al. Predicting the effective mechanical property of heterogeneous materials by image based modeling and deep learning. *Computer Methods in Applied Mechanics and Engineering*, 2019, vol. 347, pp. 735–753.
- [10] Lu W.C. et al. Using support vector machine for materials design. *Advances in Manufacturing*, 2013, vol. 1, pp. 151–159.
- [11] Song K. et al. A steel property optimization model based on the XGBoost algorithm and improved PSO. *Computational Materials Science*, 2020, vol. 174, art. no. 109472.
- [12] Delchevalerie V. et al. When magnetron sputtering deposition meets machine learning: Application to process anomaly detection. *Surface and Coatings Technology*, 2024, vol. 477, art. no. 130301.
- [13] Nazari A.G., Mozafari M. Simulation of structural features on mechanochemical synthesis of Al<sub>2</sub>O<sub>3</sub>-TiB<sub>2</sub> nanocomposite by optimized artificial neural network. *Advanced Powder Technology*, 2012, vol. 23, no. 2, pp. 220–227.
- [14] Lu W.C. et al. Using support vector machine for materials design. *Advances in Manufacturing*, 2013, vol. 1, pp. 151–159.
- [15] Chen C.T., Gu G.X. Generative deep neural networks for inverse materials design using back-propagation and active learning. *Advanced Science*, 2020, vol. 7, no. 5, art. no. 1902607.