

«Сейфуллин окулары – 18: « Жастар және ғылым – болашаққа көзқарас» халықаралық ғылыми -практикалық конференция материалдары = Материалы международной научно-практической конференции «Сейфуллинские чтения – 18: « Молодежь и наука – взгляд в будущее» - 2022.- Т.1, Ч.IV. - С. 46-48

ПРИМЕНЕНИЕ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ ДЛЯ ОПРЕДЕЛЕНИЯ СОСТОЯНИЯ ЛОКАЛЬНОЙ СИСТЕМЫ НА ОСНОВЕ ДАННЫХ ДИСТАНЦИОННОГО ЗОНИРОВАНИЯ

Елубаев Д.Р.

Казахский агротехнический университет имени С. Сейфуллина

e-mail: dastan.yelubayev@kazatu.kz

I. Введение

На сегодняшний день модели машинного обучения нашли широкое применение в рамках междисциплинарного подхода к изучению состояния развития локальных систем сельскохозяйственных посевов в купе с непрерывным мониторингом земной поверхности на основе обработки спутниковых данных дистанционного зондирования [1]. Оценка состояния системы велась непосредственно с помощью технологии Change Detection (Обнаружения изменений) трёх спектральных индексов NDVI (Нормализованный вегетационный индекс) и NDMI (Нормализованный разностный индекс влажности), NDRE (Нормализованный относительный красный краевой индекс).

II. Основная часть

В данной работе, в качестве объектов исследования использовались снимки космических аппаратов Sentinel, которые были получены благодаря облачной платформе для геопространственного анализа данных в планетарных масштабах с открытым каталогом баз данных дистанционного зондирования и исходным кодом – Google Earth Engine [2] на языке программирования Python.

Исследуемая область отображена в виде 132 полигонов (60 x 60 метров каждый) на экспериментальном поле (Рисунок 1), находящимся на территории Северо-Казахстанской области, ТОО «СХОС» - 54° 9'14.71"Сс.ш. и 69°30'34.80"В.д., система посевов – подсолнечник, пшеница, тритикале, горох, лён, просо.



Рисунок 1. Область интересов исследования экспериментальных полей.

В качестве инструментария, при помощи которого проводилось исследование, были выбраны пакеты прикладных программ с открытым исходным кодом Google Earth Engine, библиотека машинного обучения Python scikit-learn, принципиальная схема исследования представлена взаимосвязью работ библиотек геопространственного анализа данных с



библиотеками машинного обучения на рисунке 2.

Рисунок 2. Функциональная схема задачи обнаружения изменений

Данные были получены и дешифрованы за период вегетационного процесса сельскохозяйственных культур и представили собой табличные данные (Таблица 1) со значениями трёх сельскохозяйственных индексов NDVI, NDMI, NDRE, которые в последующем были подготовлены к методу кластеризации данных.

Таблица 1. Парсинг спектральных кривых яркости и их расчет индексов.

Культура	NDVI	NDVI	NDVI	NDMI	NDMI	NDMI	NDRE	NDRE
	25.05.2021	09.06.2021	16.06.2021	09.06.2021	01.07.2021	21.07.2021	31.07.2021	13.08.2021
Подсолнечник	0.114	0.156	0.192	0.249	0.408	0.733	0.762	0.667
Подсолнечник	0.113	0.157	0.232	0.269	0.408	0.758	0.776	0.695
Просо	0.119	0.157	0.232	0.289	0.424	0.757	0.777	0.694
Тритикале	0.112	0.152	0.186	0.230	0.423	0.759	0.785	0.690

Лен	0.114	0.151	0.210	0.261	0.419	0.783	0.792	0.721
.....
.....
.....
Горох	0.117	0.149	0.260	0.326	0.555	0.759	0.738	0.663

Поиск оптимального количества кластеров производился «методом локтя» в автономном обучении и определил 4 кластера (Рисунок 3):

```
X = np.array(df.loc[:,['NDVI','NDRE',
'NDMI',]]).reshape(-1, 2)
wcss = []
for i in range(1, 11):
    model = KMeans(n_clusters = i,
                    init = 'k-means++',
# инициализация метода К-средних
                    max_iter = 300,
# Максимальное число итераций
                    n_init = 10,
                    random_state = 0)
    model.fit(X)
```

```
wcss.append(model.inertia_)
# Визуализация данных
plt.plot(range(1, 11), wcss)
plt.title('Метод Локтя Elbow Method')
plt.xlabel('Количество кластеров')
plt.ylabel('Сумма квадратов внутри кластера Within Cluster Sum of Squares (WCSS)')
plt.show()
```

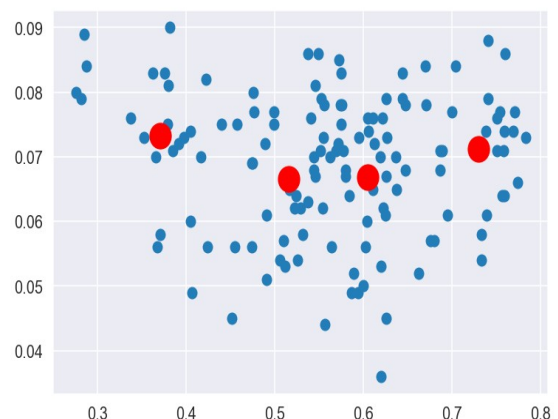
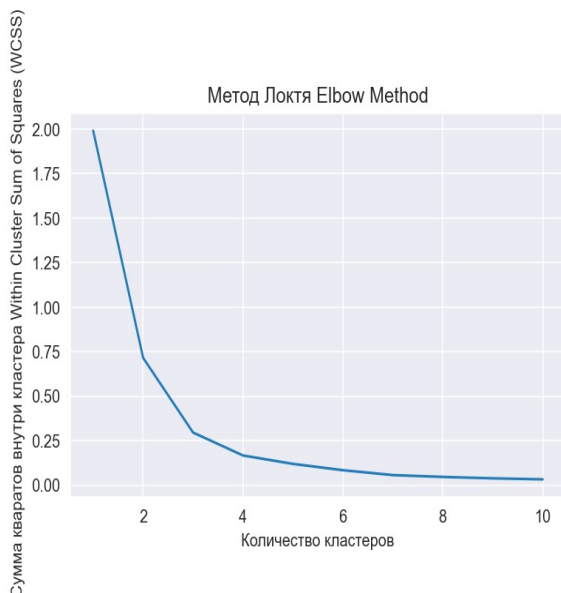


Рисунок 3. Визуализация кластеризации данных.

III. Заключение

Из-за различных условий съемки (время года, время суток, погодные условия и др.) использование традиционных методов обнаружения

изменений на изображениях не всегда эффективно. Ввиду того, что система посевов включала в себя синонимичные подсемейства сельскохозяйственных культур, модель машинного обучения K-средних лишь кластеризовала 4 класса вместо 6, что в свою очередь говорит о необходимости добавления в подготовленный банк данных еще большее количество спектральных индексов. Результаты эксперимента позволяют сформулировать требования к параметрам алгоритма предварительной кластеризации в зависимости от решаемой задачи для обнаружения изменений.

References

S. D. Suchi, A. Menon, A. Malik, J. Hu and J. Gao, "Crop Identification Based on Remote Sensing Data using Machine Learning Approaches for Fresno County, California," 2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService), 2021, pp. 115-124, doi: 10.1109/BigDataService52369.2021.00019.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone.