Goldschmidt 2026 Abstract
https://doi.org/10.46427/gold2026.33461

Delineating hydrocarbon microseepage zones in the Tarapur Block, Cambay Basin, India, using unsupervised clustering techniques

NIRAV RAJENDRA MANDORA[1,2,3], SHREY PANDYA[4], PARAS SOLANKI[2], P. L. SRINIVASA RAO[5] AND HOWARD EWAN[3]

[1]Gujarat University
[2]M. G. Science Institute
[3]S&P Global Inc
[4]The Maharaja Sayajirao University of Baroda
[5]Gujarat Energy Research and Management Institute

The Cambay Basin is a highly prospective petroliferous basin with multi-generative depressions and source rocks at various stages of maturity. Although the Tarapur Block, located in the southern-central portion of the basin, possesses all petroleum system components and features thick source-rock units, hydrocarbon discoveries have been found only along its margins. In these kinds of mature basins, conventional exploration methods combined with quantitative data-driven unsupervised machine learning techniques can help mitigate risk and increase the chance of success.

This research uses unsupervised machine learning techniques to find areas in the Tarapur Block where hydrocarbons might be leaking by analyzing surface geochemical data. In the absence of priori labeled seepage information, unsupervised classification provides an effective exploratory framework to objectively identify patterns and associations within multivariate geochemical data. Three complementary clustering techniques-Fuzzy k-means, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Gaussian Mixture Models (GMM)—were utilized to delineate both stable and transitional geochemical domains and to evaluate the robustness of the identified clusters. The clustering results show that there are geochemical patterns that are consistent over space. These patterns are interpreted to be surface signs of subsurface pathways from which hydrocarbons are seeping. Fuzzy k-means shows transitional zones with mixed geochemical signatures, HDBSCAN finds high-density anomaly clusters while filtering out background noise, and GMM gives probabilistic estimates of cluster membership. Combining these methods enables us to draw better objective lines around areas that are likely to have microseepage than any one method alone.

The study shows that the unsupervised clustering technique of surface geochemical data is a safe and cost-effective approach to delineating areas of hydrocarbon charge, in turn helping de-risk prospects and near-field and infill opportunities in more mature provinces.