

SANOAT TARMOQLARINI OPTIMAL SEGMENTATSIYALASHDA GMM VA K-MEANS USULLARINING GIBRID YONDASHUVI

DSc, dots. **Raximov Anvar**

Iqtisodiyot va pedagogika universiteti

ORCID: 0009-0009-2855-4084

sanjar.8548@mail.ru

Annotatsiya. Maqolada sanoat tarmoqlarini optimallashtirishda segmentatsiyalash hozirgi kundagi dolzarb mavzulardan biri ekanligi, hamda segmentatsiyalashda GMM modeli va K-meansning analitik sharhlari keng yoritib berilgan. Segmentatsiyalashda K-Means va GMM usullarining taqqoslama sharhi to'liq tahlil qilingan. GMM va K-means usullarini umumlashtirishning "Oqim sxemasi" yo'nalishida "Yangi GMM usuli" algoritmi va segmentatsiyalashning takomillashtirilgan gibrid modeli (HSM)ni amalga oshirish algoritmi ishlab chiqilgan. Sanoat tarmoqlarini optimal segmentatsiyalashning gibrid yondashuvi bo'yicha xulosa va takliflar berilgan.

Kalit so'zlar: sanoat tarmoqlari, optimal, segmentatsiyalash, klasterlashtirish, K-means usuli, GMM (Gaussian Mixture Model) usuli, Oqim sxemasi, algoritm, iteratsiya, adaptatsiya, integratsiya, kovariatsiya, gibrid model.

ГИБРИДНЫЙ ПОДХОД МЕТОДОВ GMM И K-MEANS В ОПТИМАЛЬНОЙ СЕКМЕНТАЦИИ ПРОМЫШЛЕННЫХ ОТРАСЛЕЙ

DSc, доц. **Рахимов Анвар**

Экономико-педагогический университет

Аннотация. В статье рассматривается вопрос о том, что сегментация при оптимизации работы отраслей промышленности является одной из актуальных тем на сегодняшний день, и представлен комплексный обзор аналитических обзоров модели GMM и метода К-средних в сегментации. Проведен сравнительный анализ методов К-средних и GMM в сегментации. В направлении «Блок-схемы» обобщения методов GMM и К-средних разработаны алгоритм «Новый метод GMM» и алгоритм реализации усовершенствованной гибридной модели сегментации (HSM). Даны выводы и предложения по гибриднему подходу к оптимальной сегментации отраслей промышленности.

Ключевые слова: промышленные отрасли, оптимальный, сегментация, кластеризация, метод К-средних, метод GMM (модель смеси Гаусса), блок-схема, алгоритм, итерация, адаптация, интеграция, ковариация, гибридная модель.

HYBRID APPROACH OF GMM AND K-MEANS METHODS IN OPTIMAL SEGMENTATION OF INDUSTRIAL INDUSTRIES

DSc, assoc. prof. **Rakhimov Anvar**
University of Economics and Pedagogy

Abstract. *The article discusses the fact that segmentation in the optimization of industrial sectors is one of the current topical issues, and provides a comprehensive overview of the analytical reviews of the GMM model and K-means in segmentation. A comparative review of the K-Means and GMM methods in segmentation is fully analyzed. In the direction of the "Flowchart" of generalization of the GMM and K-means methods, the "New GMM Method" algorithm and the implementation algorithm of the Improved Hybrid Segmentation Model (HSM) are developed. Conclusions and suggestions are given on the hybrid approach to optimal segmentation of industrial sectors.*

Keywords: *industrial sectors, optimal, segmentation, clustering, K-means method, GMM (Gaussian Mixture Model) method, Flowchart, algorithm, iteration, adaptation, integration, covariance, hybrid model.*

Kirish.

Sanoat tarmoqlarini segmentatsiyalash usuli sanoatni strategik rivojlantirishda muhim ahamiyatga ega bo'lgan jarayon bo'lib, uning ahamiyati hamda muhimlik darajalari bir qancha jihatlarida namoyon bo'ladi. Segmentatsiya sanoat tarmoqlarini xususiyatlari va ehtiyojlariga qarab guruhlash asnosida individual rivojlantirish strategiyalarini ishlab chiqish imkoniyatini yaratadi. Shuningdek, iqtisodiy diversifikatsiya va barqarorlikni ta'minlashning holati asosini tarmoqdagi kuchli, o'rta, sust rivojlangan iqtisodiy munosabatlar bilan shakllanadi. Umuman olganda sanoat tarmoqlarini segmentatsiyalash yondoshuvi iqtisodiy resurslarni optimal boshqarish, maqsadli strategiyada ijro ko'lamini real amalga oshirish mumkin bo'lgan funksiyalarni ishlab chiqish, texnologik innovatsiyalarni izchil joriy etish bilan barqaror iqtisodiy rivojlanishni ta'minlashning ajralmas qismidir. Bu jarayon sanoat tarmoqlari o'rtasida aniq ixtisoslashuv va sinergiya yaratishga, shuningdek, iqtisodiyotning uzoq muddatli barqaror rivojlanishiga xizmat qiladi.

Adabiyotlar sharhi.

Zamonaviy ekonometrik manbalar sigmentatsiyalashda ikki kuchli matematik apparatning mavjudligini targ'ib etib keladi. Bularga GMM (Gaussian Mixture Model) va K-means klasterlashtirish kiradi. Jumladan, olimlar (Kanada, Dehli) Kodamana hamda Afacan (2019) larning tadqiqotlarida sanoat tarmoqlarida o'zaro chambarchas bog'langan, uyg'unlik hosil qilgan jarayonlarning aniq tasvirlarini olish, yashirin qonuniyatlarini aniqlashda sigmentatsiyalash usulini amalga oshirish GMM bilan qulaylashtirilganligini ta'kidlaydi. Sharqiy Xitoy fan va texnologiya universiteti professori Haifei (2024) hamda tadqiqotchi olim Shibo sanoat tarmog'i va undagi iqtisodiy-ijtimoiy jarayonlar o'zining ko'p modali xususiyati va murakkab fazoviy va vaqt korrelyatsiyasi bilan tavsiflanishidan kelib chiqib sigmentatsiyalashning zaruriyatini tushuntiradi. Bunda GMM ustuvorligi va K-means zaruriyati taqqoslanadi. Asosiy jihat ma'lumotlar hajmi, kalit o'zgaruvch bilan bog'liq ekanligi asoslanadi.

Tadqiqotchi Turruellas (2024) GMM modellarini qo'llash mohiyatiga bag'ishlangan ilmiy asarlarida "GMM o'z mohiyatiga ko'ra murakkab, multimodal taqsimotni har biri ma'lumotlar ichida boshqa klasterni ifodalovchi oddiy Gauss taqsimotlarining kombinatsiyasi bilan yaqinlashishi mumkin degan prinsip asosida ishlaydi" mazmunida tasnifini keltiradi.

Tadqiqotchi Antropov (2017) to'liq bo'lmagan va ishonchsiz ma'lumotlar sharoitida havo ob'ektlarini tanib olish muammolarida K-Means va G-Means algoritmlari oilalari bo'yicha segmentatsiyalashni ko'rib chiqadi. Germaniyalik olimlar Jurg va Dennis (2019) izotropik

Gausslar bilan GMM kesilgan variatsion EM yaqinliklari qo'llanilganda, K-Means (Lloyd algoritmi) maxsus holat sifatida olinishini mumkinligini ta'kidlaydi.

Hindistonlik olimlar Patel va Kushvahalarning (2020) fikricha K-Means yordamida olingan klasterlar juda mavhum ma'lumot beradi, aksincha GMM aniq foydalanish chegaralari bilan yaxshiroq guruhlashni amalga oshiradi. Gauss aralashmasi modeli K-meansga qaraganda ko'proq hisoblash vaqtiga ega bo'lsada, undan ko'proq ish yukini tavsiflash va tahlil qilish kerak bo'lganda foydalanish mumkin" deb hisoblaydi.

E'tiborli jihat tadqiqotchilar mazkur usullarning integratsiyasi bo'yicha yetarlicha matematik murakkablik mavjudligi ta'kidlaydilar. Shuningdek, suniy intellekt (SI)ning turli tarmoqlarga va tadqiqot usullarini amalga oshirish jarayonlariga kirib kelishi algoritmnlarni qayta ko'rib chiqishni taqoza etishi fonida olimlar Guledgudd, Shariff va Quadri (2024) tomonidan Leykemiya tashxisida mashinani o'rganishni yaxshilash uchun K-means, GMM, SVM usullarining qiyosiy mukammal tahlili keltiriladi. Bu tadqiqotlarning yukunida ushbu usullarni takomillashtirish faqatgina ularning integratsiyasi bilan amalga oshishi mumkinligi ta'kidlanadi.

Tadqiqot metodologiyasi.

GMM modeli hamda K-meansning analitik sharhiga to'xtalib o'tamiz. Bizda ma'lumotlar to'plami $X = \{x_1, x_2, \dots, x_N\}$ berilgan bo'lsin. Bu yerda: x_j - n -o'lchovli vektor.

GMM modelining umumiy ko'rinishi quyidagicha:

$$f(X) = \sum_{k=1}^K \pi_k \cdot N(X|\mu_k, \Sigma_k) \quad (1)$$

Bu yerda: X - ko'p o'lchovli o'zgaruvchilar vektori, K - segmentlar soni, π_k - ushbu $\sum_{k=1}^K \pi_k = 1, 0 \leq \pi_k \leq 1$ munosabatni qanoatlantiruvchi k -segmentning aralashma og'irligi, $N(X|\mu_k, \Sigma_k)$ - k -segmentning ko'p o'lchovli normal taqsimoti quyidagi tenglik bilan ifodalanadi:

$$N(X|\mu_k, \Sigma_k) = \frac{\exp(-0.5 \cdot (X - \mu_k)^T \cdot \Sigma_k^{-1} \cdot (X - \mu_k))}{(2\pi)^{d/2} \cdot |\Sigma_k|^{0.5}} \quad (2)$$

Bu yerda: μ_k - k -segmentning markazi, Σ_k - kovariatsiya matrisasi, d - o'zgaruvchilar soni.

Shu yerda ta'kidlash joizki GMM bog'liq o'zgaruvchiga ega emas, chunki model bo'yicha bajarilgan operatsiyalar natijasi segmentlarning o'z xususiyatlariga qarab ajratiladigan guruhlari hisoblanadi. $f(X)$ GMMning ehtimollik zichlik funksiyasini bildirib, uning qiymati ma'lumotlar nuqtasining GMM tomonidan modellashtirilgan taqsimotidagi ehtimoligini ifoda etadi. Model parametrlari sifatida yuqorida izohlangan segmentlar soni (K), har bir segmentning taqsimotdagi ulushi (π_k), va o'rtacha qiymati (μ_k), hamda segmentdagi o'zgaruvchilar o'rtasidagi bog'liqlikni ifodalovchi kovariatsiya matritsasi, ya'ni (Σ_k) lar ko'rib chiqiladi. GMM modeli bu parametrlar yordamida to'liq aniqlanadi.

Ushbu parametrlar tarkibiy jihatdan quyidagi munosabatlar bilan aniqlanadi:

$$\mu_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} \cdot x^{(n)}, \quad \pi_k = \frac{N_k}{N}, \quad N_k = \sum_{n=1}^N \gamma_k^{(n)} \quad (3)$$

$$\Sigma_k = \frac{1}{N_k} \sum_{n=1}^N \gamma_k^{(n)} \cdot (x^{(n)} - \mu_k)(x^{(n)} - \mu_k)^T \quad (4)$$

Bu yerda: N_k - k -segmentdagi nuqtalar soni, N - umumiy nuqta (ma'lumot)lar soni, $\gamma_k^{(n)}$ - i -nuqta k -segmentga tegishli bo'lish ehtimoli (a'zolik ehtimoli) yoki oldindan hisoblangan shartli ehtimollik bo'lib, Expectation-Maximization (EM) algoritmi yordamida hisoblanadi:

$$\gamma_k^{(n)} = f(z_k = j | x_{jk}, \pi_k, \mu_k, \Sigma_k), j = 1, 2, \dots, n \quad (5)$$

Bu erda: z_k - maxfiy komponent ko'rsatkichini yoki x_j ma'lumotlar nuqtasi uchun yashirin segmentni ifodalaydi.

GMM modelida klasterlar sonini aniqlash muhim bosqich bo'lib, optimal K qiymatini topish uchun bir nechta usul va mezonlar qo'llaniladi. Jumladan, Elbow (tirsak) hamda Silhouette Score (Siluet ko'rsatkichi) eng ko'p foydalaniladigan usullardir.

Elbow usulida K larning turli qiymatlari uchun modelning logarifmik ehtimollik qiymati $\log L$ (Log-likelihood) hisoblanadi. Hisoblash natijalarining grafikda aks etishi "tirsak nuqtasi"ni aniqlash imkoni beradi. Bu esa K_{opt} ni ifoda etadi. Modelning logarifmik ehtimollik qiymati quyidagicha aniqlanadi:

$$\log L = \sum_{j=1}^N \log \left(\sum_{k=1}^K \pi_k \cdot N(x_j | \mu_k, \Sigma_k) \right) \quad (6)$$

Bu yerda: $N(x_i | \mu_k, \Sigma_k)$ - j -kuzatuvning k -segmentdagi ehtimollik zichligi.

Elbow (tirsak) usuli bo'yicha K_{opt} ni aniqlashda AIC (Akaike Information Criterion) yoki ko'p hollarda shartning qatilyliginiga ko'ra BIC (Bayesian Information Criterion) mezonlaridan foydalaniladi. BICning kichik qiymatida optimal kuzatiladi. BIC quyidagicha hisoblanadi:

$$BIC = p \cdot \log N - 2 \log L \quad (7)$$

$$p = K \cdot (0,5 \cdot (d^2 + 3d) + 1) \quad (8)$$

Bu yerda: p - model parametrlarining umumiy soni, d - o'zgaruvchilar soni.

Siluet ko'rsatkichi $S(j)$ segmentlar ichidagi zichlik va segmentlar orasidagi farqni baholash orqali K_{opt} ni aniqlaydi. $S(i)$ qiymat quyidagicha aniqlanadi:

$$S(j) = \frac{b(j) - a(j)}{\max\{a(j), b(j)\}}, \quad S = \frac{1}{N} \sum_{j=1}^N S(j) \quad (9)$$

Bu yerda: $a(j)$ - j -kuzatuvning o'z segmentidagi o'rtacha masofasi, $b(j)$ - j -kuzatuvning eng yaqin boshqa segment bilan o'rtacha masofasi, S - o'rtacha Siluet ko'rsatkichi bo'lib, uning eng yuqori qiymati K_{opt} ga teng bo'ladi.

Tahlil va natijalar muhokamasi.

GMM modeli bo'yicha segmentatsiyalash natijalarining aniqligini ta'minlash modelni qurish bosqichida GMMni boshlang'ich qiymatlar bilan muvaffaqiyatli moslashga bevosita bog'liq bo'ladi. Buning uchun odatda K-mean usuli algoritmidan foydalaniladi (Saheed et al., 2022). Biroq K-means algoritmi barcha jihatlariga ko'ra mos keladi deb bo'lmaydi. Aynan mana shu jihat va shu kabilar GMM modellashtirish zaruriyatini keltirib chiqaradi. Bu ikki apparatning taqqoslamasi sanoat tarmoqlarini strategik rivojlantirishda segmentatsiyalash masalasining yechimini aniqlash bilan bog'liq ularning ustuvorlik holatlarini baholashga imkon beradi. Ma'lumotlar diagonal, elliptik yoki asimmetrik joylashgan bo'lsa, GMM model K-meansga nisbatan samaraliroq. U har bir segmentni Gaussian taqsimot yordamida ifodalaydi hamda segmentlarning shakli, kattaligi va yo'nalishini modelga kiritish imkoniyatiga ega. Shunday bo'lsada, K-means usulini tadbiiq etishning o'ziga xos xususiyatlari ham mavjud bo'lib, bu ayniqsa segmentatsiyalash jarayonida yuqori o'lchamli, katta hajmdagi va yaxshi masshtablangan ma'lumotlar uchun mos keladi (Baid et al., 2017). K-means algoritmi segmentlar ichidagi ma'lumot nuqtalari orasidagi masofani minimallashtirish va segmentlar orasidagi farqni oshirishga harakat qiladi. Umumiy holda minimallashtirish tamolida ishlaydi. Minimallashtirish sharti (maqsad funksiya) Evklid masofasining kvadrati bo'yicha quyidagicha:

$$M = \sum_{k=1}^K \sum_{x_j \in C_k} \|x_j - \mu_k\|^2 \rightarrow \min \quad (10)$$

Bu yerda: C_k – k -segmentning barcha nuqtalari to'plami bo'lib, to'plam quyidagicha aniqlangan:

$$C_k = \{x_j: \|x_j - \mu_k\|^2 \leq \|x_j - \mu_j\|^2, j = 1, 2, \dots, K\} \quad (11)$$

K-meansda har bir segment uchun markaz quyidagi formula bilan yangilanadi.

$$\mu_k = \frac{1}{|C_k|} \sum_{x_j \in C_k} x_j \quad (12)$$

Bu yerda: $|C_k|$ – k -segmentning nuqtalari soni bo'lib $N_k = |C_k|$.

Segmentatsiyalash usullari bo'yicha yuqorida keltirilgan asosiy jihatlardan kelib chiqib quyidagi taqqoslash jadvalini keltiramiz (1-jadval).

1-jadval

Segmentatsiyalashda K-Means va GMM usullarining taqqoslama sharhi

T.r	Xususiyatlar	K-Means	GMM
1	Asosiy yondashuv	Segmentatsiya masofaga asoslanadi	Segmentatsiya ehtimollik ko'rsatkichiga asoslanadi
2	Yaqinlashish tamoyili	Evklid metrikasi	Gauss taqsimoti
3	Shakl moslashuvchanligi	Guruhlarining sferik shakli	Guruhlarining ellipsoid shakli
4	Hisoblash murakkabligi	Darajalangan o'lchamlarga ega emas, sodda	Darajalangan o'lchamlar mavjud bo'lganda murakkab
5	O'lchamdagi samaradorlik	Yuqori o'lchovlarda tezkor	Yuqori o'lchovlarda murakkab
6	Kovariatsiyani hisoblash	Kovariatsiyani hisobga olmaydi	Kovariatsiyani hisobga oladi

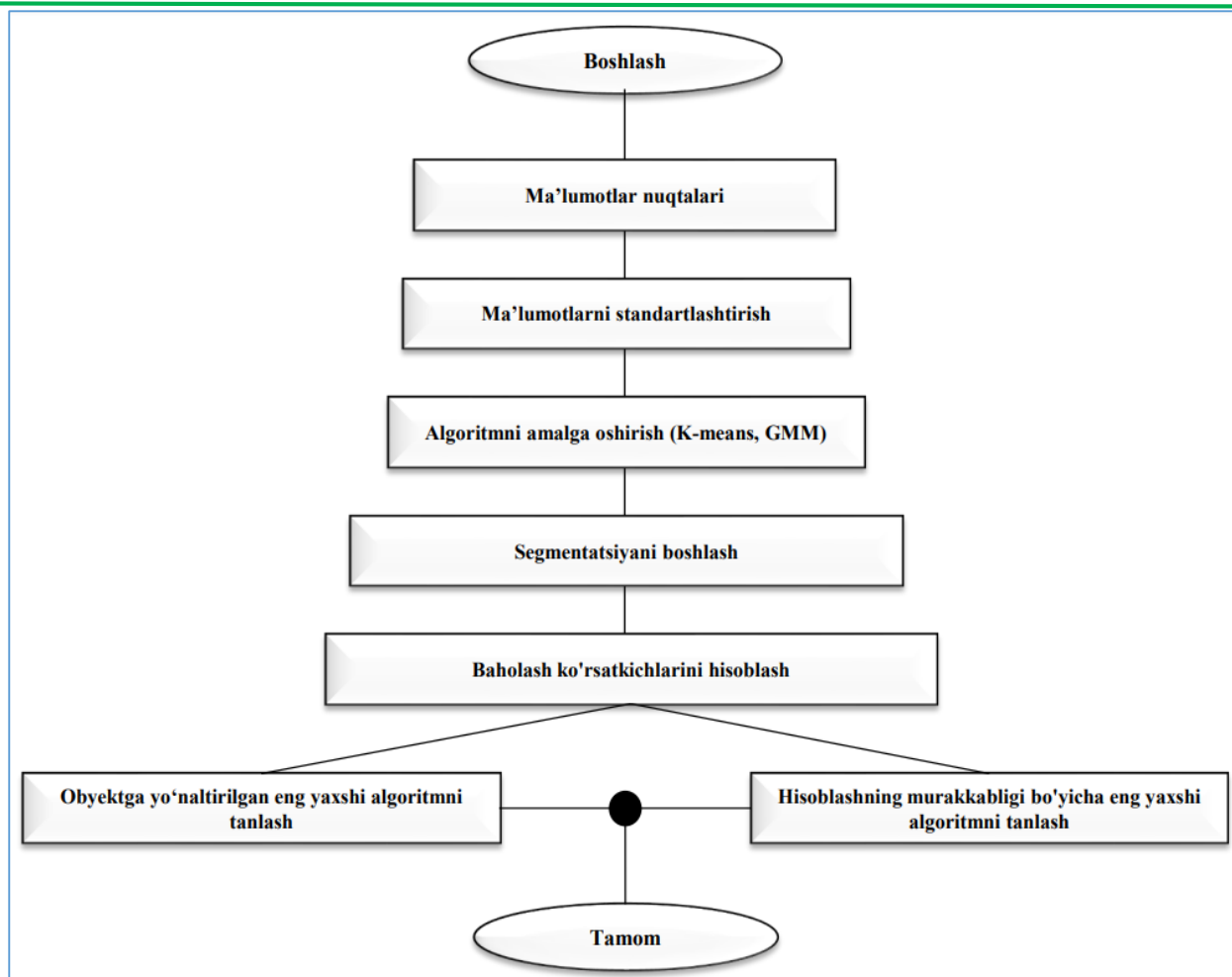
Manba: muallif tomonidan tuzilgan.

Bugungi kunda segmentatsiyalash usullarining ma'lum bir obekt yoki jarayon uchun bir vaqtda qo'llanilishi bo'yicha amaliyotlar ko'p uchraydi. Bu bo'yicha o'rganilgan manbalar shuni ko'rsatadiki, birinchidan obekt va jarayonning murakkablik xususiyatlari segmentatsiyalash natijalarining yagona usul bilan ishonchli bo'lmasligi, ikkinchidan, metodik yondoshuvlarning bir vaqtli realizatsiya qilish imkoniyati har doim mavjud bo'lishi, va provardida natijalar aniqligini oishirishga erishiladi.

Bugungu kun zamonaviy iqtisodiy modellashtirish nazariyasida GMM va K-means algoritmlarini umumlashtirishning "Oqim sxemasi" yo'nalishi rivojlanmoqda. Bu bo'yicha olib borilgan tadqiqotlarni xulosalaganda mazkur algoritmi "Yangi GMM usuli" algoritmi nomi bilan o'zlashtirilmoqda va kuchli ilmiy usulni amalga oshirish uchun qo'llanilmoqda. Quyida "Yangi GMM usuli" algoritmini ko'rib chiqamiz (1-rasm).

Ushbu algoritmi GMM va K-means usullarini amalga oshirish algoritmiga asoslansada, ko'proq GMM usuli elementlarining ketma-ket munosabatlarini yoritadi. Shuningdek, GMM va K-means usullarini umumlashtirishning "Oqim sxemasi" yo'nalishida "Yangi GMM usuli" algoritmi faqat asosiy segmentatsiyalashning bloklar tartibini aks ettiradi. Iroqlik texnolog olimlar Iman, Jawad Al-Nayar va Ali Mahmood (2023) tomonidan GMM va K-means usullarining birlashtirilgan algoritmi ishlab chiqiladi. Bu algoritmi har bir usul algoritmining birlashtirish nuqtalarini aks ettirgani holda tarmoqlanishsiz, sodda blokli tamoyillar bilan cheklanadi.

Nazariyachi A.A.Iman cheklangan energiya manbalariga ega jarayonlarda energiya sarfini minimallashtirish yoki an'anaviy texnikani yaxshilash bo'yicha yangi strategiyalarni ishlab chiqishda yangi Gauss aralashmasi modellari (GMM) algoritmini aynan 1-rasmda aks etgani kabi taklif etadi. GMM algoritmi bilan segmenlash jarayonining samaradorligini baholash tarmoqda 92% gacha energiya tejashni ko'rsatadi.



**1- rasm. GMM va K-means usullarini umumlashtirishning “Oqim sxemasi”
yo’nalishida “Yangi GMM usuli” algoritmi**

Manba: <https://beei.org/index.php/EEI/article/view/5707> /Bulletin of Electrical Engineering and Informatics Vol. 12, No. 6, December 2023, pp. 3727~3735.

Mazkur va boshqa manbalar tahlilidan kelib chiqilsa, GMM va K-means usullarining integratsion algoritmini qurish nazariy jihatdan to’liq asoslanadi. Biroq o’tkazilgan tadqiqotlar segmentatsiyalash usullarining alohida qo’llanilishi va natijalarning taqqoslamasiga asoslangan aniqlashtirilganligi ko’proq kuzatiladi. Integratsion yondoshuvlar mavjud bo’lsada aniq algoritmik va formal modeli mavjud emas. Mazkur holatda, tadqiqotimiz jarayonida segmentatsiyalashning mazkur usullarida ustunlik jihatlarini birlashtirish bo’yicha metodologik yondoshuvlar ko’rib chiqildi.

Dastlab integratsiyalashgan usulning asosiy kontseptsiyasini yaratamiz. Bunda quyidagi bir ma’noli ikki jihatga e’tibor beramiz, ya’ni

1-jihat. K-Means: Ma’lumotlarni tezkor va sodda segmentatsiya qilishda qo’llaniladi. Bu usul Evklid masofasiga asoslangan va segment markazlarini iterativ aniqlaydi;

2-jihat. GMM: Segmentlarning statistik xususiyatlarini aniqlaydi, har bir segmentni Gauss taqsimoti orqali tavsiflaydi.

Integratsiya jarayonida bu ikki bir ma’noli jihatlar K-Means usulidan boshlang’ich segment markazlarini aniqlash va GMM yordamida segmentlar parametrlarini (o’rtacha, dispersiya va ehtimolliklar) yanada aniqroq baholash bilan birlashtiriladi. Bu yerda yadro ikki qimli munosabatlardan tashkil topadi. 1-qismda (10), (11) va (12) tenglamalar, 2-qismda Gauss taqsimoti (2) tenglama, Gauss aralashma ehtimoli (1) tenglama hamda maksimal ehtimollikning ushbu

$$L(\mu_k, \Sigma_k) = \prod_{j=1}^N f(x_j | \mu_k, \Sigma_k) \quad (13)$$

tenglamasidan foydalaniladi.

Segmentatsiyalashning integratsiyalashgan usuli bizning yondoshuvimizga ko'ra quyidagi besh bosqichli algoritmgaga asoslanadi.

1-bosqich. Segemntatsiya markazlarini aniqlash (boshlang'ich K-Means segmentlash): o'zlashtirilgan qiymat - $\mu_k^{(0)}$. Tenglama: (12);

2-bosqich. Adaptatsiya jarayonini o'tkazish (GMMga moslashtirish): o'zlashtirilgan qiymat(lar) - $\Sigma_k^{(0)}, \pi_k^{(0)}$. Tenglama: (3) va (4)

3-bosqich. Expectation-Maximization (EM) iteratsiyasi (a'zolik ehtimolini ko'rib chiqish). Tenglama: E-qadam uchun (a'zolik ehtimoli)

$$\gamma_{jk} = \frac{\pi_k \cdot f(x_j | \tau_k)}{\sum_{i=1}^K \pi_i \cdot f(x_j | \tau_i)}, j = 1, 2, \dots, n, i = 1, 2, \dots, K \quad (14)$$

Tenglama: M-qadam uchun (parametrlarni yangilash)

$$\mu_k = \frac{\sum_{j=1}^n \gamma_{jk} \cdot x_j}{\sum_{j=1}^n \gamma_{jk}}, \pi_k = \frac{\sum_{j=1}^n \gamma_{jk}}{n}, \Sigma_k = \frac{\sum_{j=1}^n \gamma_{jk} \cdot (x_j - \mu_k)(x_j - \mu_k)^T}{\sum_{j=1}^n \gamma_{jk}} \quad (15)$$

4-bosqich. Parametrlarni yangilash intervalini belgilash sharti (iteratsion sikl davomiyligi). Iteratsiya jarayonida maqsad funksiyaning konvergensiya kuzatish munosabati (maksimal ehtimollik farqlari, t -iteratsiya qadami, eps – yetarlicha kichik musbat miqdor shartli tanlanadi):

$$|L^{(t+1)} - L^{(t)}| < eps \quad (16)$$

5-bosqich. Segmentatsiyani olish: Har bir nuqta uchun eng katta γ_{jk} ga mos keluvchi segment aniqlanadi.

2-jadval

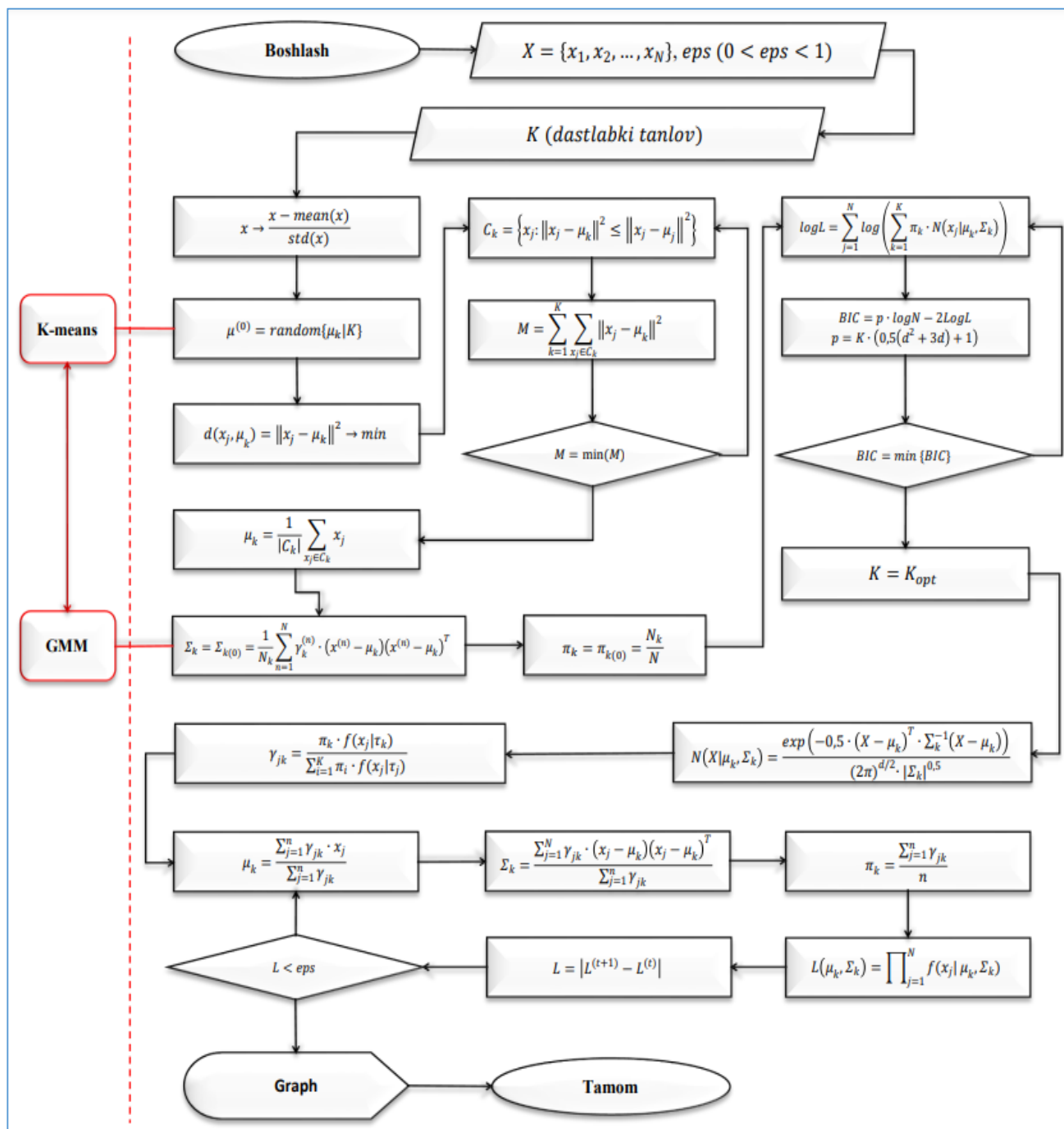
Segmentatsiyalashda K-Means, GMM va HSM usullarining taqqoslama sharhi

Xususiyatlar	Segmentatsiyalash usullari		
	K-Means	GMM	HSM
Asosiy yondashuv	Segmentatsiya masofaga asoslanadi	Segmentatsiya ehtimollik ko'rsatkichiga asoslanadi	Segmentatsiya metrik yaqin to'plamlar ehtimollik ko'rsatkichiga asoslanadi
Yaqinlashish tamoiili	Evklid metrikasi	Gauss taqsimoti	Gauus taqsimoti (Boshlang'ich nuqtalarda Evklid metrikasi)
Shakl moslashuvchanligi	Guruhlarning ellipsoid shakli	Guruhlarning sferik shakli	Guruhlarning sferik shakli
Hisoblash murakkabligi	Darajalangan o'lchamlarga ega emas, sodda	Darajalangan o'lchamlar mavjud bo'lganda murakkab	Darajalangan o'lchamlarga ega emas, sodda
O'lchamdagi samaradorlik	Yuqori o'lchovlarda tezkor	Yuqori o'lchovlarda murakkab	Yuqori o'lchovlarda tezkor
Kovariatsiyani hisoblash	Kovariatsiyani hisobga olmaydi	Kovariatsiyani hisobga oladi	Kovariatsiyani hisobga oladi

Manba: muallif tomonidan tuzilgan.

Shunday qilib bizda GMM va K-means usullarining integratsiyasi asosida takomillashtirilgan gibril modeli HSM (Hybrid segmentation model)ni taklif etish imkoniyati yaratildi. Ushbu modelning natijaviyligi K-means usuliga nisbatan GMM modeli natijalariga ko'proq yaqin ekanligini ta'kidlash o'rinli bo'ladi. Haqiqiyan, HSM usulida segment markazlari K-means markazlarini o'zlashtiradi va aniq boshlang'ich qiymatlarga ega bo'ladi.

E'tiborli jihat shundaki, HSM usuli GMM bilan bir xil natijalarni takrorlashi faqat dastlabki markazlar ustma-ust tushganda sodir bo'ladi. GMM ni amalga oshirish algoritmiga ko'ra, boshlang'ich markazlar Gauss taqsimoti natijalariga asoslanadi. Bu yerda yuqori darajalnmagan to'plam belgilarida K-means usuli markazlari bilan bir xillikni ta'minlaydi. Quyida takomillashgan HSM usulining boshqa usullar bilan xususiyatlardagi farqlar taqqoslangan (2-jadval).



2-rasm. Segmentatsiyalashning takomillashtirilgan gibril modeli (HSM)ni amalga oshirish algoritmi

Manba: muallif tomonidan tuzilgan.

Tadqiqotimiz jarayonida HMS ni amalga oshirish algoritmi ishlab chiqildi (2-rasm). Ushbu algoritm K-means va GMM algoritmlari umumlashmasidan iborat bo'lib, boshlang'ich optimal markazlarni aniqlashda to'liq K-means metrik yaqinlashishiga tayanadi. Bu esa segmentlarning tarqalish markazini qo'lay aniqlash imkonini beradi. Qolgan keyingi jarayonlar GMM usuli algoritmi asosida boradi. Algoritmida uchta sikl ishtirok etadi. Hususan K-means va GMM algoritmlarida 2 ta sikl mavjud bo'lib mazkur holatda K ning optimal ko'rsatkichi aniqlash (1-sikl), segment markazlarini generatsiyalash usulida o'rnatish (2-sikl) hamda segmentatsiya aniqligini (to'g'ri bajarilganligini) tekshirish (3-sikl) bajariladi. Aynan mana shi jihatlarni mantiqiy izchillikda birlashtirish HSM algoritmini tashkil etadi. Bizning holda taklif etilayotga segmentatsiyalashning gibrud usuli sanoatning turli tarmoqlarida rivojlanish darajalarini, zaif halqalarini umumlashtirish asosida kompleks yondoshishning optimal variantini tadqim etadi.

Shu yerda ta'kidlash o'rinliki, hisoblash texnikasi tarqqiyoti va klassik modellashirish usullarining birlashuvi bugungi kunda gibrud usullarni ishlab chiqish va amaliyotga joriy etish imkoniyatini yaratmoqda. Jumladan, HSM usulini birorta segmentatsiyalashga ixtisoslashtirilgan dasturiy paketlar bilan amalga oshirib bo'lmasada, obyektga yo'naltirilgan algoritmik tillar yordamida bunga erishish mumkin. Hususan, Phyton, R-Studio kabi dasturlash apparatlari shular jumlasidan. Ushbu paketlar gibrud usullar algoritmlarini yechishda foydalanuvchiga qulay interfeys yaratib beradi.

Xulosa va takliflar.

Xulosa qilib aytganda, sanoat tarmoqlarini rivojlantirish jarayonlari ko'plab omillarga bog'liq bo'lib, bu jarayonning o'ta murakkab holatda nomoyon bo'lishiga olib keladi. Jarayonda turli makroiqtisodiy va mikroiqtisodiy omillar ishtirok etadi va ularning har biri ma'lum ko'lamda bevosita va bilvosita ta'sir qiladi. Bu yerda o'zaro bog'liqlik xususiyatlari katta ahamiyat kasb etadi. Ushbu bog'liqlikni chuqur tahlil qilish, samarali rivojlanish strategiyalarini ishlab chiqish, istiqbolda rivojlanish tendensiyalarni omilli baholash hamda prognozlashtirish uchun ekonometrik modellashirish uslubiyoti muhim rol o'ynaydi. Ekonometrik usullar kuchli matematik-iqtisodiy apparat hisoblanib, statistik kuzatish va tahlil qilish imkonini yaratish bilan bir vaqtda sanoat tizimidagi statik qonuniyatlarni, dinamik o'zgarishlarini matematik model orqali baholash va bashorat qilish imkonini ham ta'minlaydi.

Sanoat tarmoqlarini strategik rivojlantirishda segmentatsiyalash usuli muhim ahamiyatga ega bo'lgan jarayon bo'lib, uning ahamiyati hamda muhimlik darajalari bir qancha jihatlarda namoyon bo'ladi. Bu esa, GMM va K-means usullarining integratsiyasi asosida takomillashtirilgan gibrud modeli HSM (Hybrid segmentation model)ni taklif etish imkoniyati yaratildi. Natijada HMS ni amalga oshirish algoritmi ishlab chiqildi. Ushbu algoritm K-means va GMM algoritmlari umumlashmasidan iborat bo'lib, boshlang'ich optimal markazlarni aniqlashda to'liq K-means metrik yaqinlashishiga tayanadi. Bu esa segmentlarning tarqalish markazini qo'lay aniqlash imkonini beradi. Qolgan keyingi jarayonlar GMM usuli algoritmi asosida boradi.

Adabiyotlar/Jumepamypa/Reference:

Antropov V.V. (2017) Application of k-means and g-means clustering algorithms in objects recognition. *International Research Journal*, No. 07 (61), Part 3. – p-6-9.

Baid U. et al. (2017) Comparative Study of K-means, Gaussian Mixture Model, Fuzzy C-means algorithms for Brain Tumor Segmentation. *Atlantis Press*. Vol. 137, – pp-592-597.

Eva Patel et al. (2020) Clustering Cloud Workloads: K-Means vs Gaussian Mixture Model. *Procedia Computer Science* 171. 158–167.

Guledgudd T. et al. (2024) A Comparative Study of K-Means, GMM, SVM, and Random Forest for Enhancing Machine Learning in Leukemia Diagnosis. *Afr. J. Biomed. Res.* Vol. 27(4s) (November); 4257-4268.

Haifei Peng et al. (2024) Multi-modal hybrid modeling strategy based on Gaussian Mixture Variational Autoencoder and spatial-temporal attention: Application to industrial process

prediction. *Chemometrics and Intelligent Laboratory Systems*. Volume 244, 105029 / <https://doi.org/10.1016/j.chemolab.2023.105029>.

Iman A.A. et al. (2023) A comparative study of Gaussian mixture algorithm and K-means algorithm for efficient energy clustering in MWSN. *Bulletin of Electrical Engineering and Informatics*, Vol. 12, No. 6, December, pp. 3727-3735, DOI: 10.11591/eei.v12i6.5707.

Jurg L., Dennis F. (2019) k-means as a variational EM approximation of Gaussian mixture models. *Pattern Recognition Letters*.vol-5. -p-1-14. <https://doi.org/10.1016/j.patrec.2019.04.001>.

Kodamana Zh.H., Afacan A. (2019) A GMM-MRF Based Image Segmentation approach for Interface Level Estimation. *IFAC-PapersOnLine* Volume 52, Issue 1, Pages 28-33 / <https://doi.org/10.1016/j.ifacol.2019.06.061>.

Рахимов А., Эшонқулов Ж. (2022) Саноат корхоналарида ишлаб чиқаришнинг иқтисодий қувватини ошириш механизми // *Iqtisodiyot va ta'lim*. – Т. 23. – №. 6. – С. 390-395.

Rakhimov A., Ravshanova M., Alieva M. (2023) Econometric analysis of increasing efficiency of industrial enterprises // *E3S Web of Conferences*. – EDP Sciences, – Т. 458. – С. 04004.

Raximov A.N. (2024) Sanoat tarmoqlarining milliy iqtisodiyot barqarorligini ta'milashdagi roli va rivojlantirish siyosati // *Spanish Journal of Innovation and Integrity*. – С. 150-156.

Saheed Y.K. et al. (2022) An Efficient Hybridization of K-Means and Genetic Algorithm Based on Support Vector Machine for Cyber Intrusion Detection System. *International Journal on Electrical Engineering and Informatics* - Volume 14, Number 2, June. – pp-426-442 / <https://www.ijeei.org/docs-33491964362d2da315e476.pdf>.

Turruellas J.C. (2024) Understanding Gaussian Mixture Models: A Comprehensive Guide. *Building & Telling Stories*. <https://www.linkedin.com/in/juancolamendy>.