Л.Н. Гумилев атындағы Еуразия ұлттық университетi

Ақпараттық технологиялар факультеті

«Ақпараттық жүйелер» кафедрасы

**«Интеллектуалды жүйелердің математикалық негіздері» пәні бойынша**

**Практикалық жұмыс**

**Орындаған:** Шекербек Ержан

**Топ:** АЖ-35

**Тексерген:** профессорЖукабаева Т.К

**Астана, 2023**

**Тақырыбы:** Машиналық оқытудағы AdaBoost алгоритмі AdaBoost **Практикалық жұмыстың мақсаты:** AdaBoost туралы мәлімет беру жіне мысал келтіру.

**Жұмыс орындалу аралығы:** Ағымдағы оқу жылының семестрінің 8-шы және 9-ші апталары.

**Жабдық:** Дербес компьютер

**Бағдарламалық жасақтама:** Python IDLE ортасы. Microsoft Word 2019-мәтіндік редактор.

**ТЕОРИЯЛЫҚ БӨЛІМ**

**AdaBoost (Adaptive Boosting**) - қиын үлгілерге бейімдеу арқылы әлсіз модельдердің (көбінесе шешетін ағаштардың) өнімділігін жақсарту үшін қолданылатын машиналық оқыту алгоритмі. Бұл алгоритм жіктеу және регрессия проблемаларында кеңінен қолданылады.

**Жұмыс істеу принципі:**

1. Әлсіз үлгілер: AdaBoost деректерді жіктеу үшін жеткілікті дәл болмауы мүмкін негізгі жіктеуіштерді ("әлсіз модельдер" деп аталатын) жасаудан басталады.
2. Іріктемелі таразылар: Басында оқу жиынындағы әрбір үлгіге бірдей салмақ беріледі.
3. Негізгі модельдік оқыту: AdaBoost іріктемелі таразыларға негізделген оқу жиынтығында әлсіз модельдерді жаттықтыруда. Бұл модельдердің алдыңғы модельдер дұрыс жіктелмеген үлгілерге баса назар аударатынын білдіреді.
4. Қателерді есептеу: Әрбір әлсіз модель оқытылғаннан кейін оның өнімділігі өлшенген қателерге сүйене отырып бағаланады. Аз қате жіберетін модельдер көбірек салмақ алады.
5. Үлгітаразыларын жаңарту: Үлгі таразылары дұрыс жіктелмеген үлгілердің салмағын арттыру үшін жаңартылады. Бұл жаңа әлсіз модельдерге жіктеу қиын үлгілерге назар аударуға мүмкіндік береді.
6. Үлгіні агрегаттау: AdaBoost әлсіз модельдерді бір мықты үлгіге біріктіреді. Әрбір модельдің салмағы оның соңғы үлгіде қаншалықты маңызды екенін анықтайды.
7. Final Model: AdaBoost барлық негізгі модельдердің болжамдарын олардың таразыларына сүйене отырып біріктіретін қорытынды үлгі жасайды.

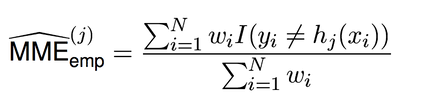
**AdaBoost артықшылықтары:**

1. Жоғары дәлдік: AdaBoost әлсіз негізгі жіктеуіштерге сүйене отырып, үлгінің дәлдігін арттыра алады.
2. Артық толтыруға төзімділік: Алгоритм үлгілердің салмағын бақылайды, бұл артық толтыру қаупін төмендетуге көмектеседі.
3. Жан-жақтылығы: AdaBoost екілік және көп сыныпты жіктеуді қоса алғанда, әр түрлі тапсырмалар түрлеріне қолданылуы мүмкін.

**AdaBoost кемшіліктері:**

1. Шудың сезімталдығы: AdaBoost деректердегі аутлер мен шуылға сезімтал болуы мүмкін.
2. Есептеу күрделілігі: Көптеген әлсіз модельдерді құру есептеу жағынан қымбат болуы мүмкін.
3. Әрдайым жарамды емес: Кейбір тапсырмаларда, егер базалық жіктеуіштер тым әлсіз болса, AdaBoost жақсы нәтижелер бермеуі мүмкін.

Формальды есептеу формуласы былай сипатталады:



Σ = қосындысы

y\_i қате жіктелген жағдайда h\_j = 1-ге тең емес және дұрыс жіктелген жағдайда 0 w\_i = салмағы

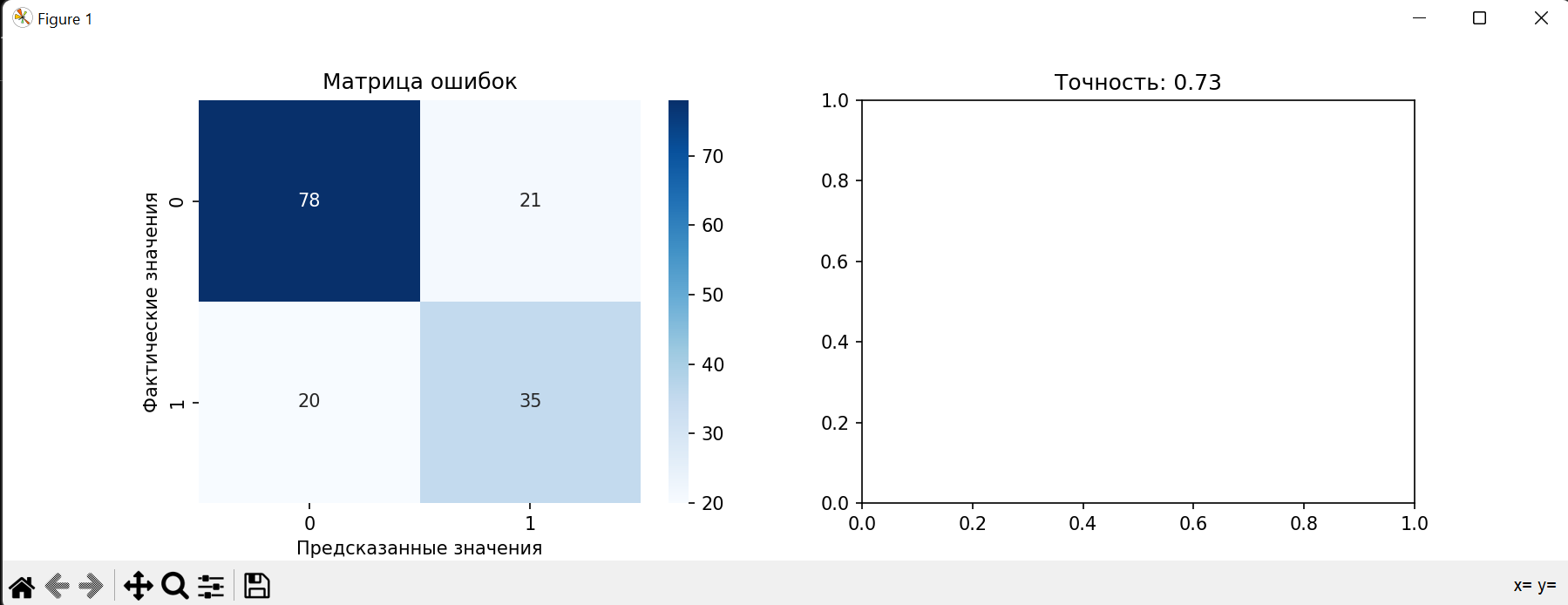
Осылайша формулада былай делінген: «Қателік қате жіктеу коэффициентінің қосындысына тең, мұндағы i және y\_i оқу үлгісі үшін салмақ біздің болжамымызға h\_j тең емес (бұл қате жіктелген жағдайда 1, ал дұрыс жіктелген жағдайда 0)».

**ПРАКТИҚАДЫҚ БӨЛІМ**

**Adaboost** моделіне қатысты шығарылған есеп келесі түрде болады:

**OUTPUT:**

**Точность модели AdaBoost: 0.7337662337662337**

**

AdaBoost моделінің дәлдігі шамамен 73,38% құрайды, бұл оның қант диабеті тапсырмасында деректерді айтарлықтай жақсы жіктеу қабілетін көрсетуі мүмкін. Алайда, модельдің болашақ болжамдарға әсерін неғұрлым терең талдау үшін бірнеше аспектілерді қарастыру керек. Егер оқу үлгісі толық болмаса немесе онда қателер болса, онда үлгі де дұрыс болмауы мүмкін. Деректерді және оның сапасын егжей-тегжейлі талдау модельдің жұмыс қабілеттілігін арттырудың негізгі факторы болуы мүмкін. Мысалы, егер медициналық диагностика үшін қант диабетін болжау маңызды болса, онда модельді науқастарда қант диабеті қаупін ерте анықтау үшін пайдалануға болады. Бұл жағдайда модель алдын алу шараларын қабылдауға және емді уақытында қабылдауға көмектесе алады. Машиналық оқыту моделi деректердiң өзгеруiне сезiмтал болуы мүмкiн және оны үнемi бақылап, жаңартып отыру қажет болады. Бұл қант диабетiнiң таралуындағы немесе медициналық практикадағы өзгерiстердi есепке алудың маңызы зор.

**ҚОЫРЫТЫНДЫ**

Осылайша, AdaBoost моделі қант диабеті сынына пайдалы болуы мүмкін, бірақ оның өнімділігі мен болашақ болжамдарға әсері тапсырманың контекстіне, деректер сапасына, өнімділік көрсеткіштеріне, сондай-ақ одан әрі оңтайландыруға байланысты болады. Барлық аспектілерге назар аударып, қажеттілікке қарай үлгіні жетілдіруді жалғастыру маңызды.