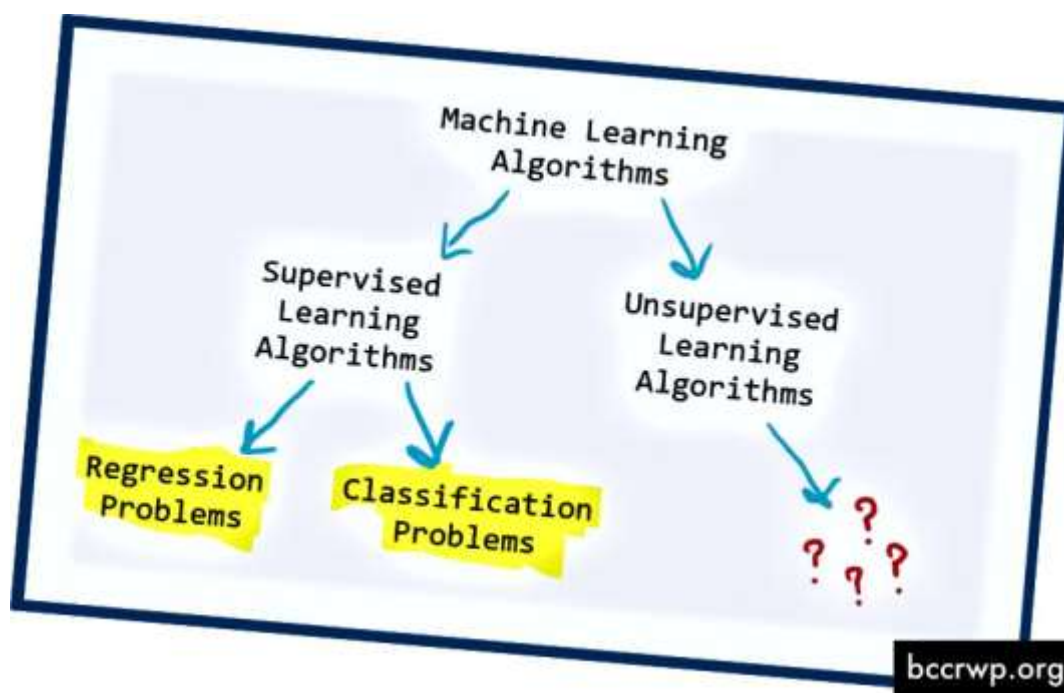


Лекция 5. Машиналық оқытудың логикалық моделі. Ранжирлеу шежіресі

1. **Лекция мақсаты мен міндеттері:** Машиналық оқытудың логикалық моделі, шешімдер шежіресінің негізгі принциптері, әдістері мен міндеттері оқыту.
2. **Лекция мазмұны:**
 1. Машиналық оқытудың логикалық модельдері.
 2. Шешімдер шежіресі.
3. **Оқытудың техникалық құралдары:** Zoom онлайн оқыту платформасы
4. **Лекция оқудың тәртібі, оқыту әдістері мен түрлері:** баяндау, сұрақ – жауап, түсіндіру, кіріспе лекция

Машинаны оқыту негіздері: сызықтық регрессия және логикалық регрессия



Сондықтан бірнеше ай бұрын мен Эндрю Нг үйрететін «Intro to Machine Learning» курсына баруды шештім. Бұл машиналық оқытудың негізгі тұжырымдамаларына арналған 11 апталық кіріспе және сіздің машинаңызды үйрену сапарын бастау үшін тамаша орын - сіз сұрағаныңызға қарамастан Бұл сапарға қайда барсаңыз да, сызықтық және логистикалық регрессияның айырмашылығын білу өте қажет. Шын мәнінде, мен бұл сұрақты өткен аптада бір сұхбатта кездестірдім және мен оған жақсы жауап бере алар едім. Мен сияқты болмаңыз, оны сұхбат алдында дәл алыңыз.

Supervised Learning

In supervised learning, we are given a data set and already know what our correct output should look like, having the idea that there is a relationship between the input and the output.

Supervised learning problems are categorized into "regression" and "classification" problems. In a regression problem, we are trying to predict results within a continuous output, meaning that we are trying to map input variables to some continuous function. In a classification problem, we are instead trying to predict results in a discrete output. In other words, we are trying to map input variables into discrete categories.

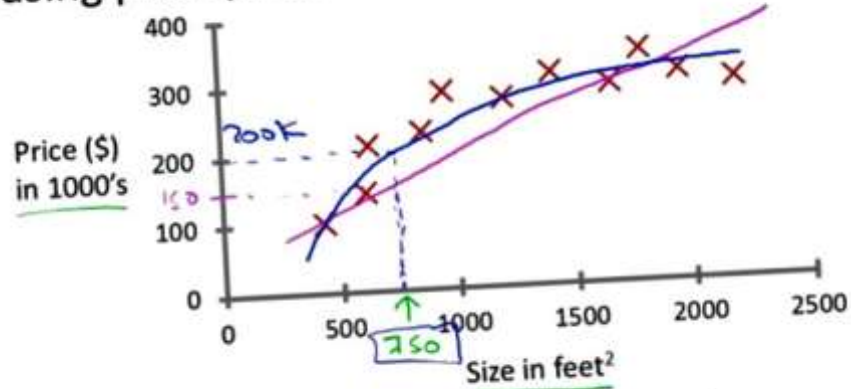
bccrwp.org

Түсінетін бірінші нәрсе, моделімізді оқытпас бұрын, жауаптың (шығудың) қандай болуы керектігін түсінген кезде, бақыланатын оқыту алгоритмдері қолданылады. Мысалы, тұрғын үй бағасын тек шаршы кадрлар үшін ғана болжау. Біздің үлгіні үйрету үшін тұрғын үй бағалары мен олармен байланысты квадраттық мәліметтер туралы мәліметтер жинауға болады, сондықтан біз бақылау алгоритмін қолдана аламыз. Егер тапсырма машинаны басқарудың орнына болса, біз мұны қалай жасау керектігін үйрету үшін мәліметтер базасын оңай құра алмаймыз. Оның орнына біз бақыланбайтын оқыту алгоритмін қолдануымыз керек еді, бірақ бұл туралы алаңдамайық.

Регрессия: сызықтық және логистикалық

Сызықтық регрессия - бұл көбінесе нүктелер жиынтығы арқылы сызық сызуға арналған модельді үйрету. Әр тармаққа жақын болған сайын, модель дәлірек болады. Бұл шамадан тыс өзгерту, бірақ көп емес! Квадрат метрлерді пайдалану үшін тұрғын үй бағасын болжау мысалын алыңыз:

Housing price prediction.



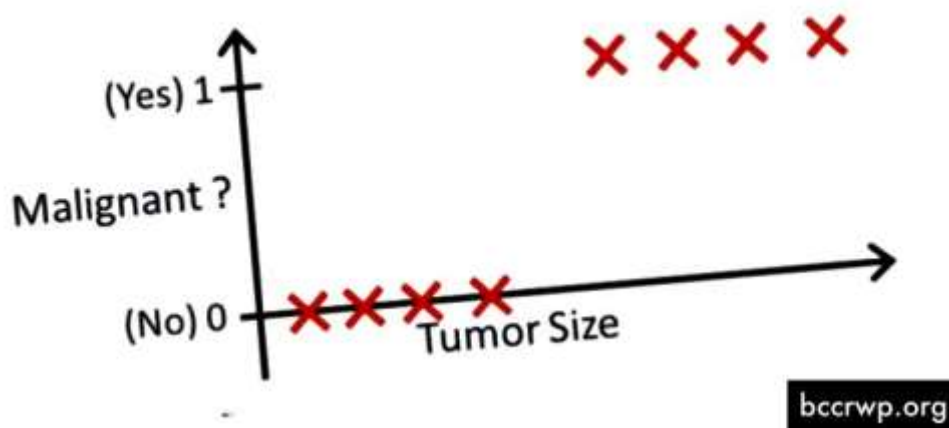
bccrwp.org

Кіріс: шаршы суреттер

Қорытынды: тұрғын үй бағасы

Жоғарыда келтірілген сызықтар сызықтық регрессияның нәтижесін көрсетеді: кіріс - квадрат, ал шығысы - тұрғын үй бағасы. Осы функцияны пайдаланып, сіз модель әлі көрмеген үйлердің тұрғын үй бағаларын нақты болжай аласыз. Егер алгоритм тек бұрын көрген үйлердің бағасын болжай алса, бұл өте пайдалы болмас еді, иә?

Логистикалық регрессия ұқсас, бірақ бірдей емес. Логистикалық регрессияда функция сызықтық емес, логистикалық. Бұл біздің нәтиже 0 мен 1 аралығында екенін білдіреді және әдетте ықтималдық ретінде түсіндіріледі (шарт бойынша). Науқастарда қатерлі ісік ауруын анықтайтын мысалды алыңыз. Шаршы суреттердің орнына біз ісіктің мөлшерін біздің нұсқа ретінде қолданамыз (айнымалы). Біздің мәліметтер жиынтығымызда ісіктердің мөлшері болады және олар қатерлі ісік па, жоқ па (0 немесе а 1 түрінде берілген).



Кіріс: ісік мөлшері

Шығу: қатерлі ісіктің ықтималдығы

Енді сіз «бұл классификацияға ұқсайды, мен регрессия туралы айтып отырмыз» деп ойлайтын шығарсыз, сіз дұрыс айтасыз. Логистикалық регрессияны жіктеу мәселесін шешуге тағы бір қадам жасауға болады (және ол солай болады!). Егер біз 0,5-тен жоғары ықтималдылық қатерлі ісік, ал одан аз нәрсе қатерлі ісікті білдіреді деп шешсек, онда voila! Біз үздіксіз енгізу, АКА, классификациядан дискретті категорияларды жасадық. Мен бұл семантика сияқты көрінетінін білемін, бірақ айырмашылықты білу өте маңызды!

Сонымен, Регрессия кіріс айнымалыларын сызықтық немесе логистикалық болсын, кейбір тұрақты функцияға картаға түсіру үшін қолданылады, ал классификация кіріс айнымалы мәндерін дискретті санаттарға салыстыру үшін қолданылады.

Логистикалық регрессияны жіктеу мәселелерін шешуге қолдануға болады, бірақ ол әлі де регрессия болып табылады.

Бұл түсінік сізге ұнады деп сенемін. Сіз сондай-ақ Эндрю Нгтің «Машиналарды үйрену» курсынан өттіңіз бе? Сіз оны қарастырып жатырсыз ба? Сіз машинаны үйрену сапарында қайда жүрсіз?

5. Деңгейлік тапсырмалар:

1-деңгей.

- Сызықтық регрессия дегеніміз не?
- Логикалық регрессия дегеніміз не?

2-деңгей.

- Машиналық оқытуда регрессияға анықтама беріңіз?
- Логистикалық регрессияны жіктеу мәселелері?

6. БОӨЖ тапсырмалары:

Ағаш шешімдерін күшейту.

7. БӨЖ тапсырмалары:

Новиковтың конвергенция туралы теоремасы. Новиков теоремасының дәлелі

8. Пайдаланылған әдебиеттер:

1. Ғ. Айғараева, Қ. Асанова, М. Нысанов. Сандық қондырғылар және микропроцессорлық жүйелер. Оқулық. 2010
2. Д. С. Козен. Есептеу теориясы. Сандық. Оқулық Алматы. 2013.
3. Асамбаев А.Ж. Жасанды интеллект негіздері.- Алматы: ЖШС РПБК "Дәуір", 2011-135б. Оқу құралы.
4. Рассел С., Норвиг П. Жасанды интеллект: Жаңашыл әдіс -Алматы, 2014.. Оқу құралы.
5. К. Пол. Цифровое искусство. Учебник. Москва , 2017, 220
6. Андреас Мюллер, Сара Гвидо. Введение в машинное обучение с помощью Python. –Москва. 2016.
7. В.Н.Негода. Машинно-ориентированное программирование. Учебное пособие Ульяновск 2015.