

مراحل انجام الگوریتم پر کردن داده های خالی :

- ابتدا داده را از طریق کتابخانه ی pandas می خوانیم و در قالب یک لیست چند بعدی ذخیره میکنیم
- برای حذف آلفا درصد از داده های یک ستون به صورت زیر عمل میکنیم
  - به تعداد آلفا درصد داده ها عدد تصادفی تولید میکنیم (بین ۰ تا انتهای لیست دیتاست)
  - داده های تصادفی به دست آمده ایندکس محل هایی هستند که بایستی داده حذف شود
  - به جای داده ی حذف شده عدد 1- قرار میدهیم
- پر کردن مقادیر با روش های :
  - میانگین : در ابتدا میانگین تمام داده هایی که 1- نباشند را محاسبه میکنیم و روی لیست حاوی مقادیر حذف شده قدم زده و به جای مقادیر حذف شده ، مقدار میانگین کل که قبلا حساب کرده ایم را قرار میدهیم
  - میانگین هم کلاسی ها : ابتدا یک دیکشنری از لیبل ها ایجاد میکنیم و با قدم زدن روی دیکشنری به ازای هر لیبل میانگین داده هایی که محتوی این لیبل باشند و 1- نباشند را محاسبه کرده برابر با مقدار لیبل متناظر در دیکشنری قرار میدهیم
  - Knn : با استفاده از کتابخانه ی sklearn.neighbors و متد KNeighborsRegressor در آن مقادیر خالی را پر میکنیم
- فاصله ی اقلیدسی بین داده های اصلی و پر شده به صورت مجذور جمع توان دوم اختلاف بین بردار یک بعدی در هر داده میباشد
- برای نمایش روی نمودار به این صورت عمل میکنیم که :
  - گرافی با یک ردیف و ۳ ستون در نظر میگیریم
  - محور X هر سه گراف مقادیر آلفا هستند
  - محور Y هر گراف متناظر با فاصله ی اقلیدسی به ازای آلفا در روش مربوطه