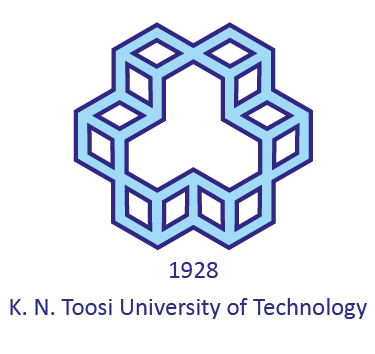
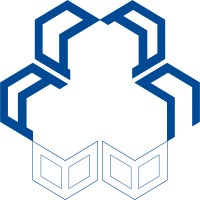
به نام خدا





دانشگاه صنعتی خواجه نصیرالدین طوسی

دانشکده برق

**شبیه سازی و مدلسازی**

**گزارش تمرین شماره 1**

**سجاد رجبی باغستان**

**40005393**

**استاد : جناب دکتر مهدی علیاری**

بهمن 1402

##### فهرست مطالب

|  |  |
| --- | --- |
| عنوان | شماره صفحه |

بخش۱: چکیده 3

[بخش۱: سوالات تحلیلی 4](#_Toc128223055)

[سوال اول 4](#_Toc128223056)

[بخش۲: سوالات شبیه سازی 5](#_Toc128223058)

[سوال اول 5](#_Toc128223059)

[سوال دوم 6](#_Toc128223060)

[سوال سوم 7](#_Toc128223060)

[مراجع 9](#_Toc128223066)

**چکیده**

در این تمرین، به بررسی و مقایسه دو روش انتخاب ویژگی Forward Selection و backward elimination میپردازیم.

Forward Selection یک روش انتخاب ویژگی‌های مهم در مدل‌های رگرسیون است. در این روش، از یک مدل اولیه شروع می‌شود که فقط یک ویژگی دارد. سپس ویژگی‌های دیگر به تدریج اضافه می‌شوند و هر بار اثر اضافه شدن یک ویژگی را ارزیابی می‌کنیم. ویژگی‌ای که اضافه شدن آن باعث بهبود بیشتری در عملکرد مدل می‌شود، به مدل اضافه می‌شود.

Backward Elimination یک روش انتخاب ویژگی در مدل‌های رگرسیون است که به تدریج ویژگی‌های غیرضروری را از مدل حذف می‌کند تا به بهترین مدل برسد. در این روش، از یک مدل اولیه که شامل تمام ویژگی‌ها است شروع می‌شود، سپس ویژگی‌های غیرضروری یکی‌یکی حذف می‌شوند و عملکرد مدل با هر حذف ارزیابی می‌شود.

مزایای این روش شامل کاهش ابعاد داده‌ها، سرعت بالا در مواقعی که تعداد ویژگی‌ها زیاد است، و قابلیت استفاده از معیارهای مختلف برای ارزیابی مدل است. از معایب آن می‌توان به این اشاره کرد که ممکن است در مواقعی ویژگی‌های مهم حذف شوند و همچنین این روش بیشتر در مواقعی مؤثر است که تأثیرات ویژگی‌ها به طور مستقل از هم نباشند.

# بخش۱: سوالات تحلیلی

## سوال اول

در بحث مدل سازی ،هدف این است که مدلی ارائه شود که نه بیش از حد بی دقت باشد و نه بیش از حد با دقت (اصطلاحا overfit یا underfit نباشد). با داشتن رگرسور های فراوان برای مدل سازی ،مدل ارائه شده به صورت خیلی دقیق عمل میکند به طوری که مدل ما overfit میشود ، همچنین با داشتن تعداد زیادی رگرسور ، مقدار زیادی نویز وارد مدل ما میشود زیرا هر رگرسور همراه با نویز است.

در ضمن این موضوع را باید در نظر داشت که ممکن است اثر یک رگرسور نسبت به بقیه رگرسور ها خیلی بیشتر باشد و همین امر باعث این میشود که استفاده از بقیه رگرسور ها ناصرفه باشد.

**دلایل ضرورت :subsection**

* به طور کلی در واقعیت تعداد رگرسور ها میتواند بسیار زیاد باشد و همین امر باعث پیچیدگی زیاد مدل از نظر محاسباتی شود به طوری که انجام این محاسبات ناممکن و یا بسیار زمان بر است.
* در بسیار مواقع سیستم تنها از تعداد محدودی رگرسور تاثیر اصلی را میپذیرد و بقیه رگرسور ها با توجه به هزینه ای که برای مدل سازی صرف میشود (زمان، محاسبات، تجهیزات،...)غیر ضروری هستند.

**فواید :subsection**

* **کاهش زمان و منابع**
* **بهتر شدن مدل و دور شدن مدل از overfitting**

# بخش۲: سوالات شبیه سازی

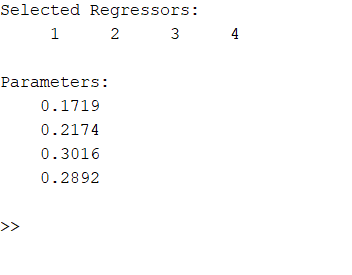
## سوال اول

در این سوال بعد از import کردن دیتا ها ، رگرسور ها را یک به یک انتخاب کرده و اضافه میکنیم

این کار را تازمانی ادامه میدهیم که کمترین تغییرات را برای ساخت y پیدا کنیم.

رگرسور هایی که انتخاب میشوند بیشترین کواریانس را با y دارند.

Result:

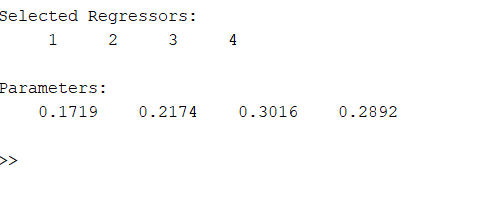


## سوال دوم

در این سوال بعد از import کردن دیتا ها ، رگرسور ها را یک به یک انتخاب کرده و حذف میکنیم

این کار را تازمانی ادامه میدهیم که بیشترین تغییرات را در y مشاهده کنیم.

Result:



## سوال سوم

|  |  |  |
| --- | --- | --- |
| ویژگی | Forward Selection | Backward Elimination |
| پایین | از تمام داده‌های گذشته استفاده می‌کند | سیستم‌های ایستا |
| بالا | وزن بیشتری به داده‌های اخیر می‌دهد | سیستم‌های پویا با سازگاری نرم |
|  |  |  |
|  |  |  |
|  |  |  |
| متوسط (ناگهانی) | داده‌های قدیمی را به‌صراحت حذف می‌کند | سیستم‌های پویا با تمرکز بر روند اخیر |

جدول1: مقایسه متد های طراحی مدل

# مراجع

Websites:

<https://chat.openai.com/>

<https://en.wikipedia.org>

<https://www.sciencedirect.com/topics/computer-science>

<https://www.simplilearn.com>

book:

Nonlinear system identification 2020 nelles