

HW2 ANN

مدی مهر
۱۴۰۱/۲۲/۱۳۹

سؤال اول:

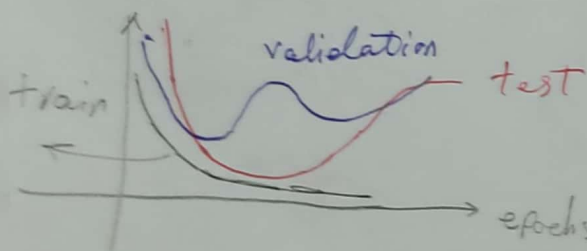
overfit :
دشمنی

الف)
یعنی اینکه شبکه من fit به روی نمونه‌هایش از اندازه
که ازل نمونه‌های representative برای کل داده نیست اگر ازل
در بیش از حد یادگیر قابلیت تعمیم خود را برای تهیه نمونه از دیت
من دست دهد. این گونه خطای شبکه ما در داده‌هایی که در طول
یادگیری اینم داده بی‌رکمی شود اما خطای شبکه برای نمونه‌ها
نست به از من رود. یعنی خطای ما در داده‌های train
پایین است است ولی در داده‌های Validation بالا است
Test

نکته: دقت در تعداد داده‌ها
ما کم باشد احتمال overfit شدن
شود ما با بالاتر من رود.
پس باز یاد گرفتن داده‌ها
رنگی که من شود.

ب) هدف به تعداد نودن‌ها بسته از حد شود مدل پیچیده‌تری شود
و امکان overfit شدن در مدل افزایش می‌یابد.
برای حل این مشکل من نویسم با نودن کم شروع می‌کنیم و برای
گرفتن حاصل بهترین تعداد نودن‌ها من ارقام ما به یک

Trade of بین حاصل خطای ما در داده‌های train
و دقت‌های Validation داریم یعنی اگر افزایش تعداد نودن ما باعث افزایش
خطای Validation شد تعداد نودن‌ها را کاهش می‌دهیم تا به یک نقطه
test
پایین برسیم.



۱- با تعداد نودن کم نویسم یعنی

۲- از $train + test$ کم‌ترین مقدار را انتخاب می‌کنیم

۳- Decay weight by small fraction

after each iteration

1) داده (ب) منطبق است داده های ما برای Train کم به همین به بیست
 overfit شدن شود برای حل این دست از موارد می توانیم بخاردها را

روش Commit
 به k دسته مختلف تقسیم کنیم از training و test دین k دسته
 به این تقسیم کنیم و به اینک اینها برای تغییر مدل استفاده کنیم.

2) می توانیم دین 4 تا نقطه را فقط بدین 4 دسته + کد را بنویسیم.

بیشتر از 2. تا نقطه 4 تا را بدین 4 دسته + کد را بنویسیم در 14

بقیه Train را این 4 دسته تقسیم دین 4 تا را وارد داده های
 آموزش می کنیم در 4 تا را به عنوان 4 دسته + کد را بنویسیم
 در Train را این 4 دسته تقسیم دین 4 تا را وارد داده های استفاده می کنیم البته
 به صورت مقطعی - افعال overfit شدن کم می گردد

نکته د باید بین پیچیدگی مدل
 با تعداد epoch تغییر
 Trade off وجود داشته باشد
 اگر تعداد epoch را دین
 مدل بیشتر از حد زیاد کنیم
 مدل ما به دردی داده و overfit
 می شود - شکل منحنی
 قبل

3) وقتی رخ می دهد که models آموزش به طور قابل ملاحظه
 training set

بیشتر از حد I deal به شد وقتی رخ می دهد که مدل قوی

model / روی مجموعه آموزش ضعیف عمل می کند و می توانیم مدل را bias

به بیستی است. که برای حل این مشکل در شبکه عصبی می توانیم

تعداد نرون ها را به دین یا تعداد لایه ها را. یک مدل under fit

شده به اندازه قوی نیست که بتواند به نظر از داده های آموزش یاد بگیرد

همچنین با افزایش دین یا دین می توانیم به این پیچیدگی

کنیم که بین صفر مدل ما به اول نقطه optim خود را

نرسیده بودن بین تعداد epoch ها را از این

بسیار تعداد epoch
 ها را تا جایی که مدل
 به حالت over fit
 نرسیده باشد می بینیم
 و همچنین اگر مدل ما
 پیچیده است تعداد لایه ها
 و نرون ها داخلی را افزایش
 می دهیم تا به اول
 حد از دست که می توانیم
 به دردی داده های مدل
 train

(نکته) برای دسته‌بندی دو کلاس، می‌توانیم از تابع فیلد استفاده کنیم

از **activation function** به معنی **فعال‌سازی** استفاده می‌کنیم. از این تابع برای پیاده‌سازی **Sigmoid** استفاده می‌کنیم.

$$\sigma(z) = \frac{1}{1 + e^{-az}}$$

استفاده می‌کنیم که شیب نزایست است و می‌توان این تابع را نیز به‌کار برد:

$$\sigma(z) = a \cdot \tanh(bz) = \frac{2a}{1 + e^{-2bz}} - a \rightarrow \tanh$$

استفاده کرد زیرا بازه پاس (۱+، ۱-) نگه می‌دارد و می‌توان در مدل قبلی بازه بین (۰، ۱) بود و در جمیع می‌توانیم در شبکه عصبی کلاسیک ممکن است صفده عنوان حاصل خروجی نداشته باشیم.

البته انتخاب این قضیه پس از هر جندی به منابع تحت اقداری ما برای می‌تواند بستگی دارد و هر قوی‌تر باشد می‌توانیم از تعداد صفده‌ها کمتر و بهتر برای پس‌بینی استفاده کنیم.

- البته از مدل‌های ساده‌تر relu نیز استفاده می‌کنیم (از آنجایی که خروجی relu برای مقدار بزرگ‌تر از صفر به‌صورت bound غیره و برای مقادیر کمتر صفر به‌صورت ثابت است و هم‌چنین ثابت بودن مشتق آن در مدل‌های پیچیده ممکن است زودتر از local minimum و در وقت آموزش زودتری به‌دست می‌آید و بسیار باهوش‌تر دارد)

برای **loss** ها **Binary Cross-entropy** می‌توانیم به‌صورت پس‌بینی در مدل **Classification** استفاده می‌کنیم.

به عنوان تابع استفاده می‌شود. به صورت ریاضی اثبات می‌کنیم که دارای **maximum** است.

Cross entropy یک معیار محاسبه می‌کند که تاکنون تفاوت بین توزیع احتمالی واقعی و پس‌بینی

شده را بداند. پس‌بینی کلاس ۱ را به‌صورت ۱ می‌دهیم و خروجی آن یک احتمال

بین ۰ و ۱ است. مثلاً اگر حاصل به ۰.۹۲ می‌باشد و احتمال ۰.۹۲ در هر

حالت ۱ در صحت است.

این برای یک شبکه عصبی خردی یک Sigmoid activation function برای لایه خروجی است -

این نادرست هم که loss را نیز یک cross entropy activation function برای لایه های دیگر است

لایه های میانی هم تابع relu را پس نادرست هم زیرا نسبت به Sigmoid در لایه های میانی برای عملکردی بهتر نقطه ای که سیر می کند است البته به سیرت سرعت (مقدار هم بستگی دارد اگر تعداد epoch ها را به زیاد کردیم ممکن نداشته باشد) در لایه های میانی نیز از Sigmoid به tanh نیز استفاده کرد

در لایه های میانی

شکل =
کلاس ه سمت چپ

سوال 2) شبکه عصبی تکوین 2 :

راه حل اول) اگر ما شکل برای ما قابل مشاهده باشد از اینجا می که تقسیم در کلاس

به دلیل یابین بودن تعداد راه های Train قابل تغذیه و سواره سازی

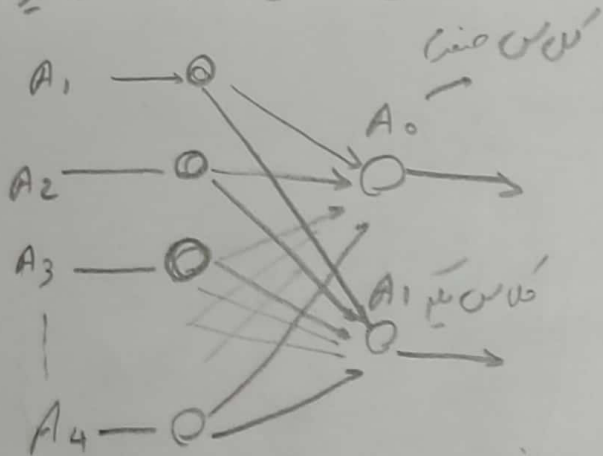
به یک boolean است که به راحتی با تعداد کم از نود و ورودی ها قابل
function
حل است.

راه حل دوم) به صورت محوس چون ما داده ها را نمی بینیم و داده ها و رابطه بین ادنها

باید توسط شبکه پیدا شود راه حل جامع تر و استوار تر می هست که در ادامه می آید.

راه حل دوم ← راه حل اصلی

ما یک مارتین داریم که ۱۲ درایه دارد که هر درایه خودش توان یک ورودی را به دو خروجی که به هم روشن یا خاموش بودن این اندیس بدست می آید



سپس یک آرایه با ۱۴ اندون

ورودی و ۲ اندون خروجی

داریم که هیچ لایه بیانی

برای شناسایی آن نداریم

برای همین فیل های خروجی از اندون های ورودی

همین و با یک one یعنی از کلاس های صفی یا یک صفی

hot cooling

با از آنجایی که احتمال شکل با بهترین شباهت به در کلاس در کلاس مربوط به خودش تکرار می شود بدین کیفیت چقدر شباهت بین این شکل و شکل های غرض آموزش وجود دارد

شکل ایالات است چپ: ۴ رنگی - ۴ سفید
شکل ایالات است راست: ۴ رنگی - ۴ سفید

توسط رنگی - ۴ سفید

ادامہ سوال 2 ب

متوسط

کارزلی - 2 سفید

→

پاسین کت چپ 4 کارزلی

2 سفید

پاسین کت راست 4 کارزلی

2 سفید

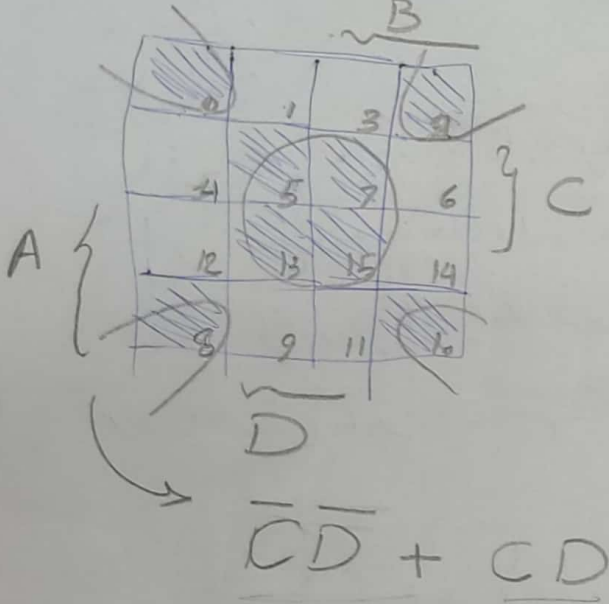
از انجایی کہ متوسط سببیت بہ فلاس ہوں

بیتہ است \Rightarrow نہ فلاس ہوں اقیماں استغفر

مجلس

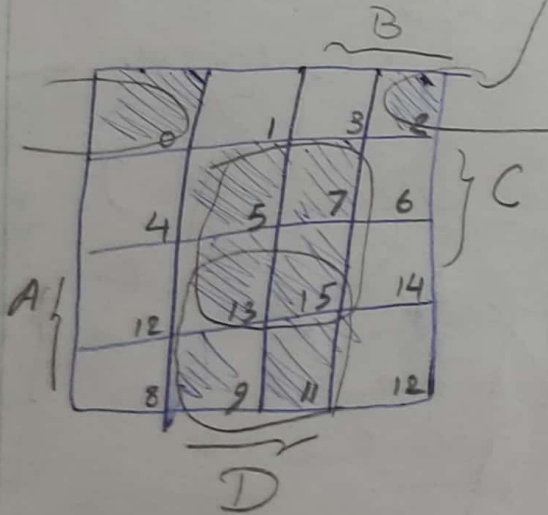
حلاسن صنفه:

سبد عصی سوال ۲)



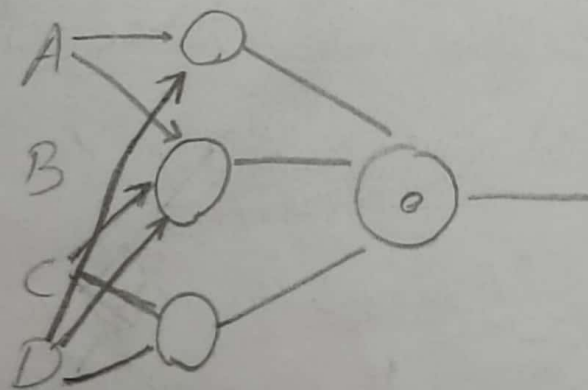
من توانم به سبد ها داده شده به عنوان یک جدول کارنو شده کنیم. که مربع های رنگ شده به فضای ۱ و مربع های سفید را صنفه در جدول میسیم و برای هر کدام از سبدها یک فرد منقضی برای آن بدست می آید که برابر است با ۱

سبد فرد اول سبت راست



$CD + AD + \overline{A}\overline{C}\overline{D}$

سبد
حال براساس این دو سبد عصی میسیم زیرا



راه حل اول

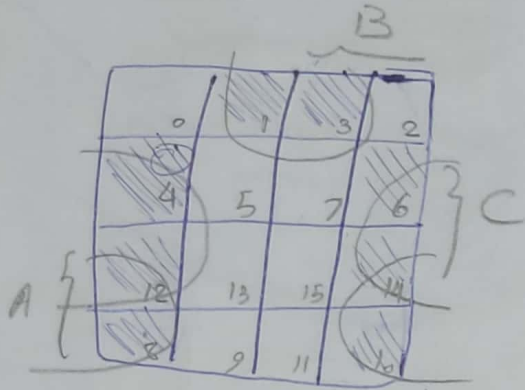
سؤال ۲ شبکۀ عصبی مخرب (۲)

محل ۱ د

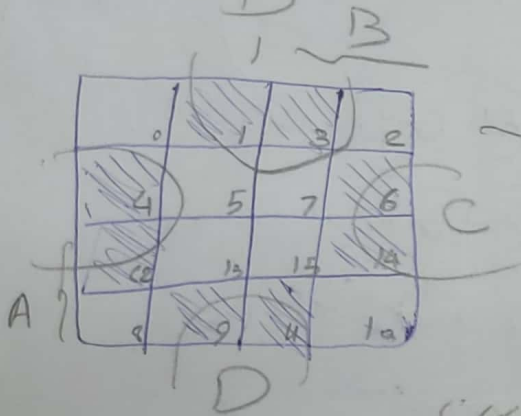
همین کار را برای شبکه های پایین

با انتخاب ۴ ورودی نیز می توانیم

فرمول منطقۀ خروجی را بدست می آوریم

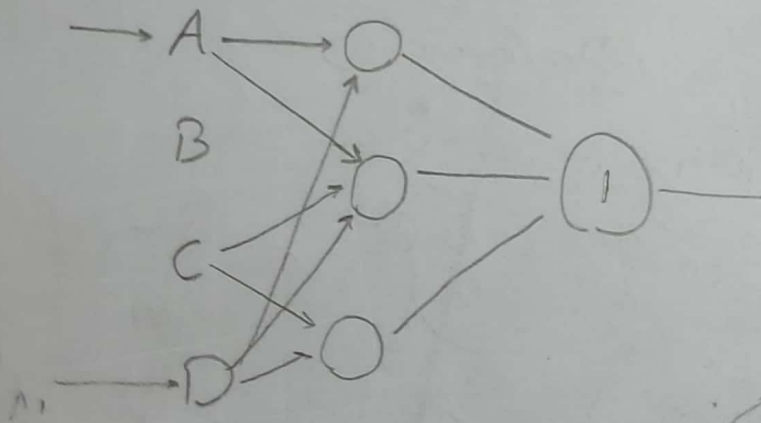


$$A\bar{D} + C\bar{D} + \bar{A}\bar{C}D$$

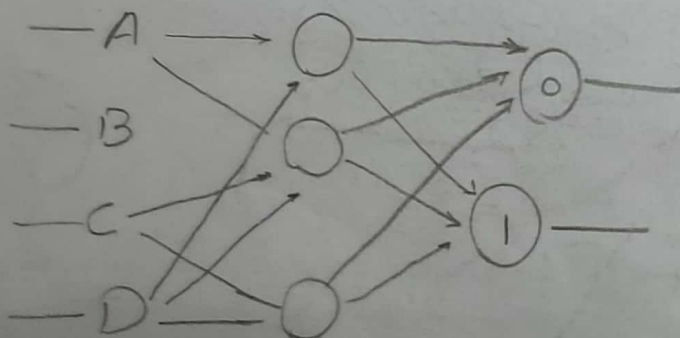


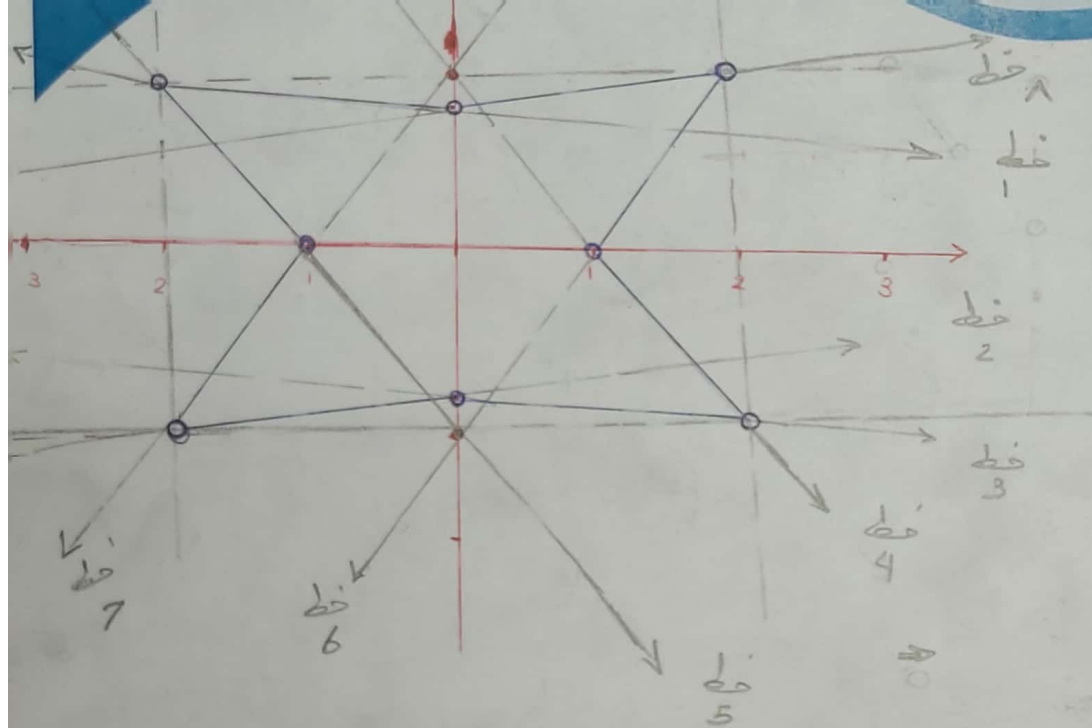
$$C\bar{D} + \bar{A}D$$

از درخت
حال بر اساس فرمول بدست آمده یک شبکۀ عصبی می بینیم



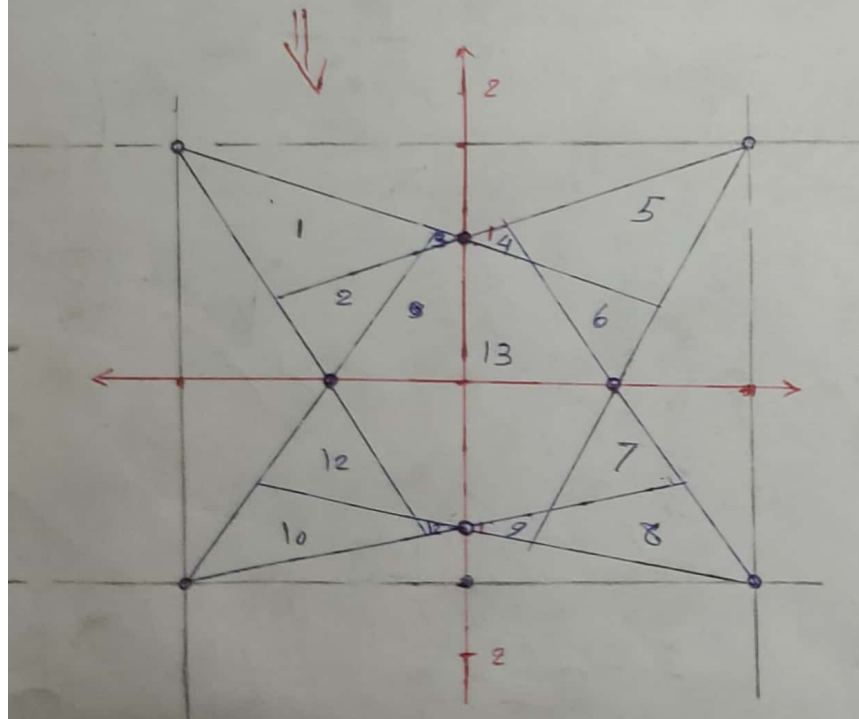
حال اگر بخواهیم حاصل را به یکدیگر ارائه کنیم





سؤال 3

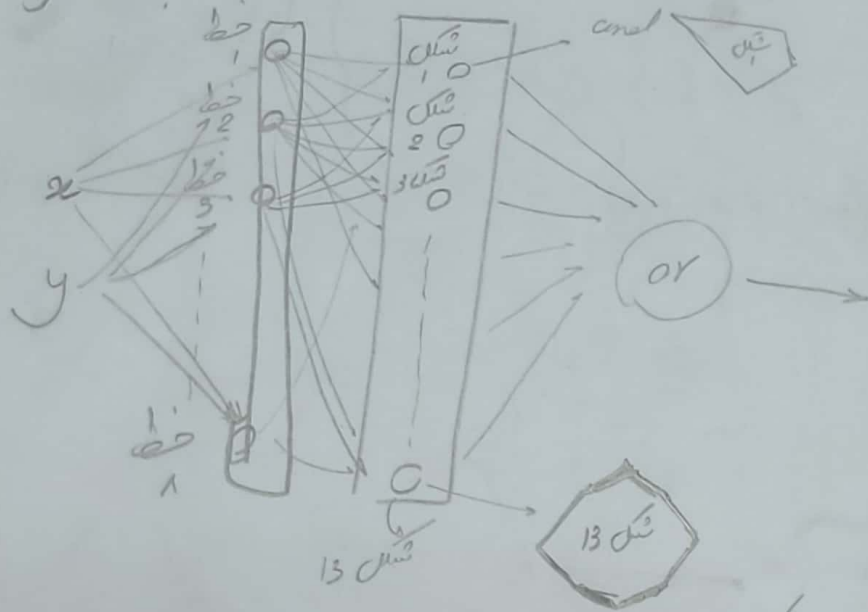
الف)



ب)

سین حاصل،
حاصل 13 سین موجب
است.

کایه میانی لایه میانی شبکه منظم:



ب لایه میانی
فید زیر اسلوت توانایی
شخص شکل ۵
خوب را داریم به نظر
برای شخص شکل
معمول مانتو شکل قبل با
کبد مانتو شبکه فون طرح نیم
حاصل با دو لایه

خب همانطور که از شکل الف شخص است ابتدا خط هایی که شکل را

تشکیل داده اند را می کشیم که شکل از ۸ خط تشکیل شده است

پس در مرحله بعد شکل هایی که از برخورد خط ها حاصل می شوند را

بدان می کشیم که ۱۳ شکل جدید به یک خط ها نامیده شده شکل ها گرد

حال اگر بخواهیم فوق به حل شده بپردازیم برای ساعت ۸ خط به

۸ جدول ورودی ۴ دل را می کشیم و می کشیم (برای ساعت

۱۳ نامیه های مورد نظر به ۱۳ نامیه می کشیم که به یک خط ها شکل ها حاصل

را به زد. حال اگر یک جدول به عنوان خودی در نظر بگیریم به نظر می آید

جمع حالت به نام می کشیم و مورد نظر قرار داریم بانه.

بله همین گفته است اگر تعداد جدول های میانی کمتر از ۱۳ تا باشد دل به دست

نی تواند حاصل را به بد زیر اگر کمتر از این تعداد باشد نامیه ها نامیه معتبر نیستند

در بعضی موارد خواهند که شکل شخص نامیه جدید به یک خط ها شکل درخته آن

شکل پذیر نیست. همین اگر تعداد خط ها کمتر از ۸ تا گردند شکل پذیر نیست

دیا اگر ہم اُنہی پذیرا ہوں، مگر کم وقت خواہد بود

پس ہیک ^{یابی} لایم بدقت مناسب نہ توانیم مگر را پیش بینا کنیم و بآگاہ نردن کم

در صورت در لایم بودن نیز طبعی اتفاق ممکن است بینند .