



CRF条件随机场讲解

分享人: grant



HMM(生成模型): $\lambda = (\pi, A, B)$

1. 齐次一阶马尔可夫假设

$$P(y|y_{1:t-1}, x_{1:t-1}) = P(y_t|y_{t-1})$$

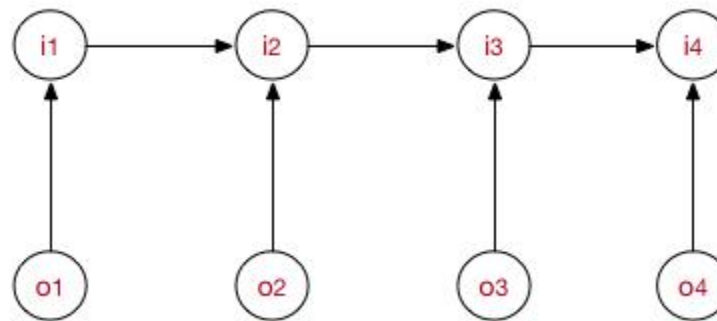
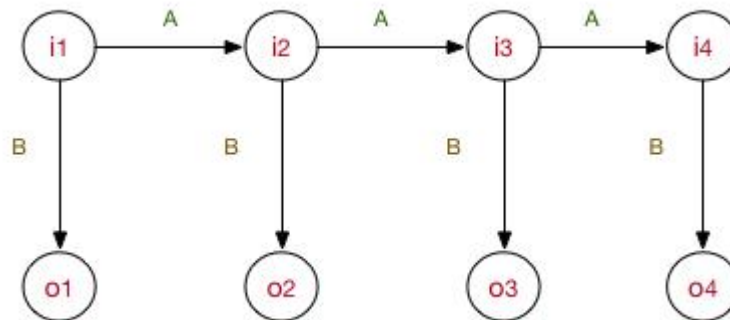
2. 观测独立假设

$$P(y|y_{1:t-1}, x_{1:t-1}) = P(y_t|y_{t-1})$$

MEMM(判别模型):

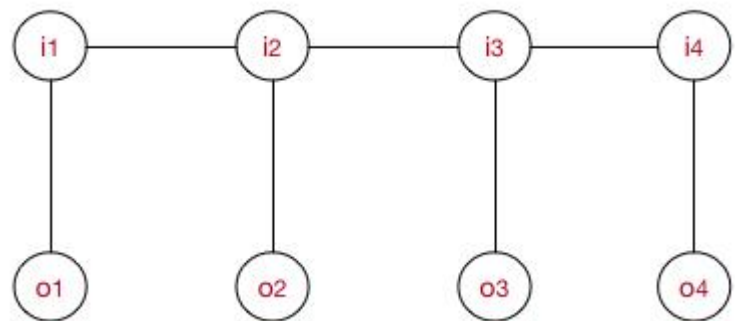
1. 打破观测独立性假设

2. 存在标记偏置的问题 (各个状态之间的转移不是随意的, 模型学到了偏好路径, 不关心观测值, 根本是因为局部归一化)



CRF(判别模型):

定义: 满足 $P(Y_v|X, Y_\omega, \omega \neq v) = P(Y_v|X, Y_\omega, \omega \sim v)$ 的马尔科夫随机场叫做条件随机场(CRF)



CRF建模公式:

$$P(I|O) = \frac{1}{Z(O)} \prod_i \psi_i(I_i|O) = \frac{1}{Z(O)} \prod_i e^{\sum_k \lambda_k f_k(O, I_{i-1}, I_i, i)} = \frac{1}{Z(O)} e^{\sum_i \sum_k \lambda_k f_k(O, I_{i-1}, I_i, i)}$$

($I_i \sim O_i$) 为最大团 $\psi_i(I_i|O)$ 为最大团上随机变量的联合概率

$Z(O)$ 为规范化因子, 对无向图所有最大团的概率进行累加

CRF的参数化形式:

$$P(I|O) = \frac{1}{Z(O)} e^{\sum_i^T \sum_k^M \lambda_k f_k(O, I_{i-1}, I_i, i)} = \frac{1}{Z(O)} e^{[\sum_i^T \sum_j^J \lambda_j t_j(O, I_{i-1}, I_i, i) + \sum_i^T \sum_l^L \mu_l s_l(O, I_i, i)]}$$

$s_l(O, I_i, i)$ $l=1, 2, \dots, L$, L 是定义在该节点的节点特征函数总个数

$t_j(O, I_{i-1}, I_i, i)$ $k=1, 2, \dots, K$, K 是定义在该节点的局部特征函数的总个数

$$s_{l=1}(o, i) = \begin{cases} 1 & \text{满足特定状态条件, 比如当前token的POS是'V',} \\ 0 & \text{other} \end{cases}$$

$$t_{k=1}(o, i) = \begin{cases} 1 & \text{满足特定转移条件, 比如前一个token是'T',} \\ 0 & \text{other} \end{cases}$$




维特比算法解码思路:

定义*i*处的局部状态为 $\delta_i(l)$, 表示在位置*i*标记*l*各个可能取值(*l*=1,2,...,*m*)对应的非规范化概率最大值, 递推在位置*i*+1标记*l*的表达式:

$$\delta_{i+1}(l) = \max_{1 \leq j \leq m} \{ \delta_i(j) + \sum_{k=1}^K w_k f_k(y_i = j, y_{i+1} = l, x, i) \}, l = 1, 2, \dots, m$$

局部状态来记录 $\Psi_{i+1}(l)$ 使 $\delta_{i+1}(l)$ 达到最大的位置*i* 的标记取值, 该值用来回溯最优解, $\Psi_{i+1}(l)$

$$\Psi_{i+1}(l) = \arg \max_{1 \leq j \leq m} \{ \delta_i(j) + \sum_{k=1}^K w_k f_k(y_i = j, y_{i+1} = l, x, i) \}, l = 1, 2, \dots, m$$


维特比算法流程:

输入: 模型的K个特征函数和对应的K个权重, 观测序列 $x = (x_1, x_2, \dots, x_n)$, 可能的标记个数m

输出: 最优标记序列 $y^* = (y_1^*, y_2^*, \dots, y_n^*)$

1) 初始化:

$$\delta_1(l) = \sum_{k=1}^K w_k f_k(y_0 = \text{start}, y_1 = l, x, i), l = 1, 2, \dots, m$$

$$\Psi_1(l) = \text{start}, l = 1, 2, \dots, m$$

2) 对于 $i = 1, 2, \dots, n - 1$, 进行递推:

$$\delta_{i+1}(l) = \max_{1 \leq j \leq m} \{ \delta_i(j) + \sum_{k=1}^K w_k f_k(y_i = j, y_{i+1} = l, x, i) \}, l = 1, 2, \dots, m$$

$$\Psi_{i+1}(l) = \arg \max_{1 \leq j \leq m} \{ \delta_i(j) + \sum_{k=1}^K w_k f_k(y_i = j, y_{i+1} = l, x, i) \}, l = 1, 2, \dots, m$$

3) 终止:

$$y_n^* = \arg \max_{1 \leq j \leq m} \delta_n(j)$$

4) 回溯:

$$y_i^* = \Psi_{i+1}(y_{i+1}^*), i = n - 1, n - 2, \dots, 1$$

最终得到最优标记序列 $y^* = (y_1^*, y_2^*, \dots, y_n^*)$

例题:

假设输入的是三个词的句子, 即 $X = (X_1, X_2, X_3)$, 输出的词性标记为 $Y = (Y_1, Y_2, Y_3)$, 其中 $Y \in \{1(\text{名词}), 2(\text{动词})\}$

这里只标记出取值为1的特征函数如下:

$$t_1 = t_1(y_{i-1} = 1, y_i = 2, x, i), i = 2, 3, \lambda_1 = 1$$

$$t_2 = t_2(y_1 = 1, y_2 = 1, x, 2) \lambda_2 = 0.6$$

$$t_3 = t_3(y_2 = 2, y_3 = 1, x, 3) \lambda_3 = 1$$

$$t_4 = t_4(y_1 = 2, y_2 = 1, x, 2) \lambda_4 = 1$$

$$t_5 = t_5(y_2 = 2, y_3 = 2, x, 3) \lambda_5 = 0.2$$

$$s_1 = s_1(y_1 = 1, x, 1) \mu_1 = 1$$

$$s_2 = s_2(y_i = 2, x, i), i = 1, 2, \mu_2 = 0.5$$

$$s_3 = s_3(y_i = 1, x, i), i = 2, 3, \mu_3 = 0.8$$

$$s_4 = s_4(y_3 = 2, x, 3) \mu_4 = 0.5$$

求标记(1,2,2)的最可能的标记序列。

解答:

首先初始化:

$$\delta_1(1) = \mu_1 s_1 = 1 \quad \delta_1(2) = \mu_2 s_2 = 0.5 \quad \Psi_1(1) = \Psi_1(2) = start$$

接下来开始递推, 先看位置2的:

$$\delta_2(1) = \max\{\delta_1(1) + t_2 \lambda_2 + \mu_3 s_3, \delta_1(2) + t_4 \lambda_4 + \mu_3 s_3\} = \max\{1 + 0.6 + 0.8, 0.5 + 1 + 0.8\} = 2.4 \quad \Psi_2(1) = 1$$

$$\delta_2(2) = \max\{\delta_1(1) + t_1 \lambda_1 + \mu_2 s_2, \delta_1(2) + \mu_2 s_2\} = \max\{1 + 1 + 0.5, 0.5 + 0.5\} = 2.5 \quad \Psi_2(2) = 1$$

再看位置3的:

$$\delta_3(1) = \max\{\delta_2(1) + \mu_3 s_3, \delta_2(2) + t_3 \lambda_3 + \mu_3 s_3\} = \max\{2.4 + 0.8, 2.5 + 1 + 0.8\} = 4.3$$

$$\Psi_3(1) = 2$$

$$\delta_3(2) = \max\{\delta_2(1) + t_1 \lambda_1 + \mu_4 s_4, \delta_2(2) + t_5 \lambda_5 + \mu_4 s_4\} = \max\{2.4 + 1 + 0.5, 2.5 + 0.2 + 0.5\} = 3.9$$

$$\Psi_3(2) = 1$$

最终得到 $y_3^* = \arg \max\{\delta_3(1), \delta_3(2)\}$, 递推回去, 得到:

$$y_2^* = \Psi_3(1) = 2 \quad y_1^* = \Psi_2(2) = 1$$

即最终的结果为(1, 2, 1), 即标记为(名词, 动词, 名词)。



谢 谢





1

年度工作概述

单击此处编辑您要的内容，建议您在展示时采用微软雅黑字体，本模版所有图形线条及其相应素材均可自由编辑、改色、替换。



目录

CONTENT

01

年度工作概述

Please replace text, click add
relevant headline, modify the text

02

工作完成情况

Please replace text, click add
relevant headline, modify the text

03

成功项目展示

Please replace text, click add
relevant headline, modify the text

04

工作评价

Please replace text, click add
relevant headline, modify the text

05

明年工作计划

Please replace text, click add
relevant headline, modify the text



项目负责人
Manager



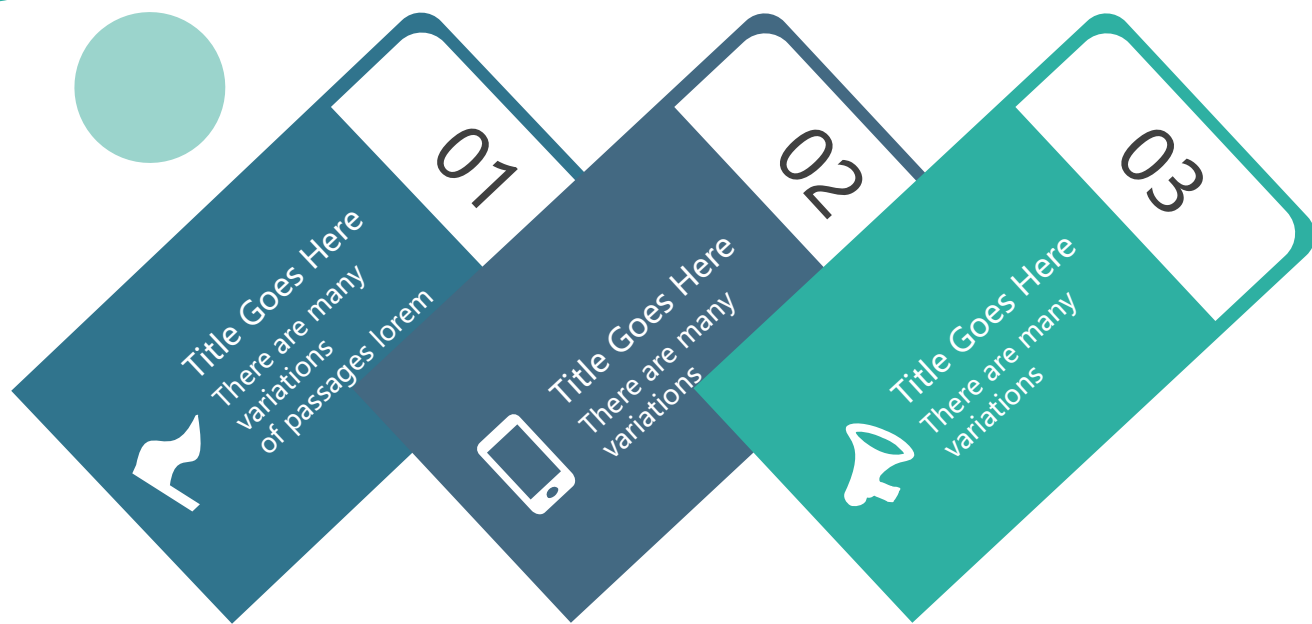
公司经理
Programmer



招商合作
Designer



工作详情
Web Designer



产品研发

Youth is not a time of life; it is a state of mind; it is not a matter of rosy cheeks, red lips and supple knees;

产品生产

Youth is not a time of life; it is a state of mind; it is not a matter of rosy cheeks, red lips and supple knees;

产品销售

Youth is not a time of life; it is a state of mind; it is not a matter of rosy cheeks, red lips and supple knees;

产品后勤

Youth is not a time of life; it is a state of mind; it is not a matter of rosy cheeks, red lips and supple knees;

项目1

There are many
variations
of passages lorem
ipsum



项目2

There are many
variations
of passages lorem
ipsum



项目3

There are many
variations
of passages lorem
ipsum



项目4

There are many
variations
of passages lorem
ipsum





2

工作完成情况

单击此处编辑您要的内容，建议您在展示时采用微软雅黑字体，本模版所有图形线条及其相应素材均可自由编辑、改色、替换。

工作成果展示

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its fulfillment regardless of obstacles. By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to



TOPIC HEADER ONE
We have many PowerPoint templates that has been specifically designed.



TOPIC HEADER TWO
We have many PowerPoint templates that has been specifically designed.





办公室工作情况

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I

have a serious reason



管理人员

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I

have a serious reason



意见反馈

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I



合作商

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I have a serious reason



社会联系

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I



市场调查

when he was a little boy ask the indulgence of the children who may read this book for dedicating it to a grown-up. I have a serious reason



产品介绍

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its fulfillment regardless of obstacles.



Information

when he was a little boy ask the
indulgence of the children who may
read this book for

20XX



年终成果汇报



Information

when he was a little boy ask the
indulgence of the children who may
read this book for

20XX



3

成功项目展示

单击此处编辑您要的内容，建议您在展示时采用微软雅黑字体，本模版所有图形线条及其相应素材均可自由编辑、改色、替换。

产品销售图

Sed in ex sed libero rhoncus molestie id vitae libero. Aenean semper posuere ullamcorper. Phasellus maximus tempus scelerisque. Nulla accumsan augue eget mauris tincidunt, vel tempus leo pharetra. Sed in velit blandit, sagittis ante sit amet, porttitor dolor. Donec egestas ultrices urna sit amet.



EDI
T



GEARS



HEARTBEAT
T





产品流程图

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its fulfillment regardless of obstacles. By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to

22.2

女性占比及原因

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its



77.8

% 男性占比及原因

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its

产品使用性别占比

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its fulfillment regardless of obstacles. By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to



4

工作评价

单击此处编辑您要的内容，建议您在展示时采用微软雅黑字体，本模版所有图形线条及其相应素材均可自由编辑、改色、替换。

工作完成情况大数据分析

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its

OUR SERVICES

Creative Project

Creative Project

Creative Project

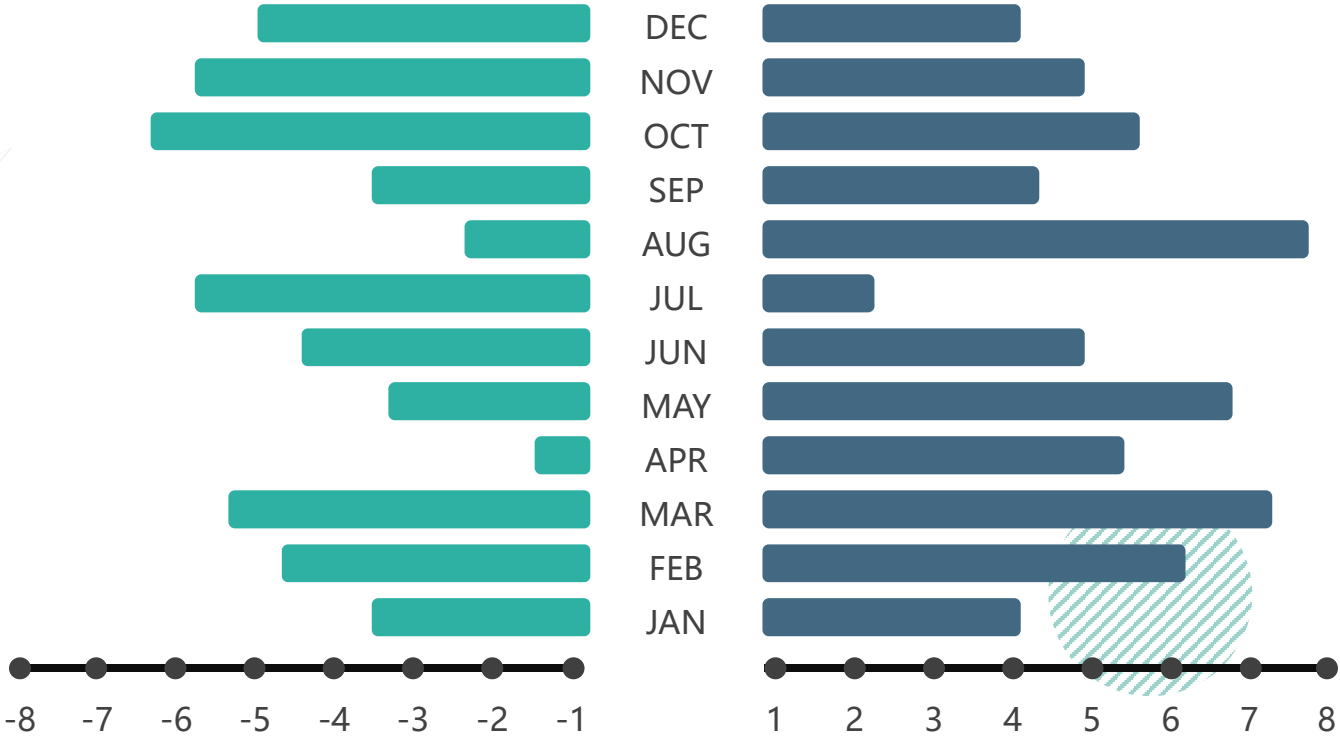
工作欠缺总结

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its

工作质量分析 相关因素

WE PROVIDE THE BEST SERVICES

By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek its fulfillment regardless of obstacles. By faith I mean a vision of good one cherishes and the enthusiasm that pushes one to seek





产品1

Design inspiration For
presentation

产品2

Design inspiration For
presentation

产品3

Design inspiration For
presentation

产品4

Design inspiration For
presentation

产品5

Design inspiration For
presentation



77.8%

同类型产品市场占比

By faith I mean a vision of good one cherishes and the enthusiasm that pushes
one to seek its





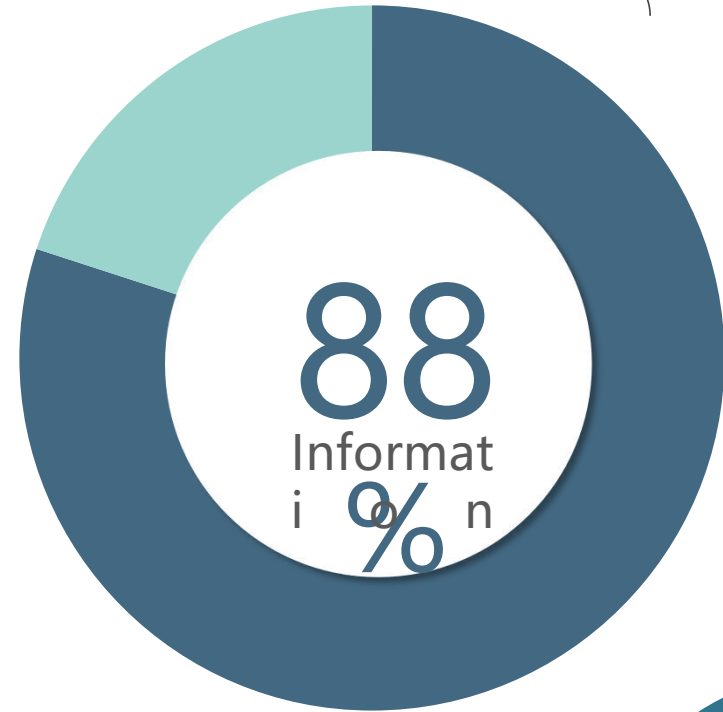
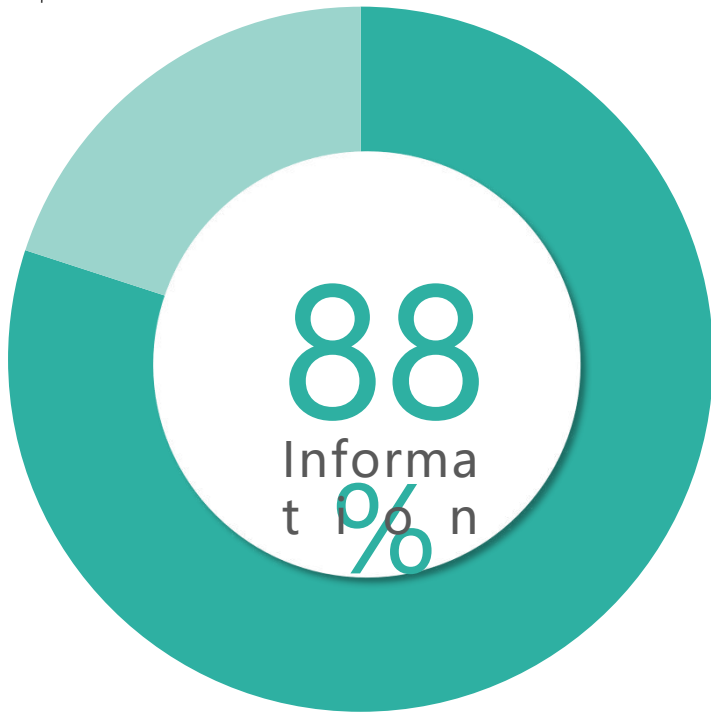
5

未来工作安排

单击此处编辑您要的内容，建议您在展示时采用微软雅黑字体，本模版所有图形线条及其相应素材均可自由编辑、改色、替换。

未来市场规划

Powerpoint offers word processing, outlining, drawing, graphing, and presentation management tools- all designed to be easy to use and learn



We are more powerful when we are clear, and we are weaker when we are confused.
Clarity Leads to Power. When we're clear, we're more effective. Our clarity reduces our mistakes and enables us to enlist the help of others.

未来产品设计



Header here please

We are placing great text in this text frames. Lorem ipsum is here mate check it out



Header here please

We are placing great text in this text frames. Lorem ipsum is here mate check it out



Header here please

We are placing great text in this text frames. Lorem ipsum is here mate check it out



Header here please

We are placing great text in this text frames. Lorem ipsum is here mate check it out

LOGO

未来项目

The past is gone and static. Nothing we can do will change it. The future is before us and dynamic.

BUSINESS



TITLE HERE

Entrepreneurial activities
differ substantially
depending on the type
of



第一PPT

HTTP://WWW.1PPT.COM

可以在下列情况使用

- 个人学习、研究。
- 拷贝模板中的内容用于其它幻灯片母版中使用。

不可以在此以下情况使用

- 任何形式的在线付费下载。
- 网络转载、线上线下传播。

PPT模板: www.1ppt.com/moban/
节日PPT模板: www.1ppt.com/jieri/
PPT背景图片: www.1ppt.com/beijing/
优秀PPT下载: www.1ppt.com/xiazai/
Word模板: www.1ppt.com/word/
个人简历: www.1ppt.com/jianli/
手抄报: www.1ppt.com/shouchaobao/
教案下载: www.1ppt.com/jiaoan/

行业PPT模板: www.1ppt.com/hangye/
PPT素材: www.1ppt.com/sucal/
PPT图表: www.1ppt.com/tubiao/
PPT教程: www.1ppt.com/powerpoint/
Excel模板: www.1ppt.com/excel/
PPT课件: www.1ppt.com/kejiao/
试题下载: www.1ppt.com/shiti/
字体下载: www.1ppt.com/ziti/