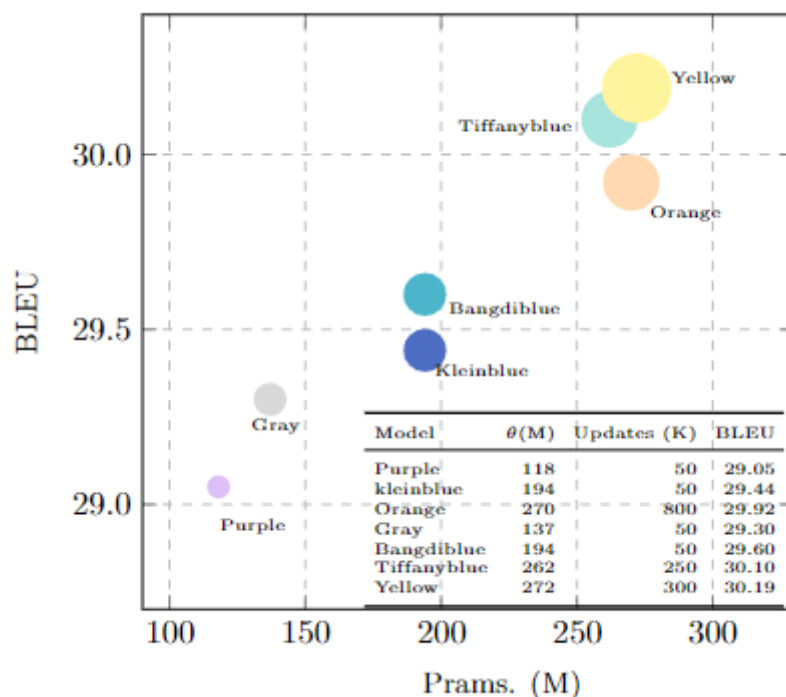


2. 如何绘制美妙的散点图



本章节主要介绍如何采用latex绘制散点图。这里主要采用的package是pgfplots。故首先要引用该包如下：

```
\usepackage{pgfplots}
```

其它需要的package同第一节一样。

pgfplots 的使用是嵌入在tikz内部的，故其代码块应该包在tikz代码块内部，如下所示：

```
\begin{tikzpicture}
...
{pgfplots 代码块;}
...
\end{tikzpicture}
```

接下来，我们将进入正题，开始从0到1绘制一个散点图。如上图所示，该散点描述了模型训练更新次数，模型容量，以及模型的性能之间的关系。

Step 1: 首先第一步我们要定义我们的画布。这一部分主要采用的是axis命令来得到我们的画布，如下所示：

```
\begin{tikzpicture}
\begin{axis}[
at={(0,0)},
ymajorgrids,
xmajorgrids,
grid style=dashed,
```

```

width=0.7*\textwidth,
height=0.65*\textwidth,
xlabel={\small{Prms. (M)}},
ylabel={\small{BLEU}},
ylabel style={yshift=0em, xshift=0em},
xlabel style={xshift=1em,yshift=0.0em},
yticklabel style={/pgf/number format/precision=1,
/pgf/number format/fixed zerofill},
ymin=28.7,ymax=30.4, ytick={29,29.5,30},
xmin=90,xmax=330,xtick={100, 150, 200, 250, 300},
]

...

\end{axis}
\end{tikzpicture}

```

其中

```
\begin{axis}[...] \end{axis}
```

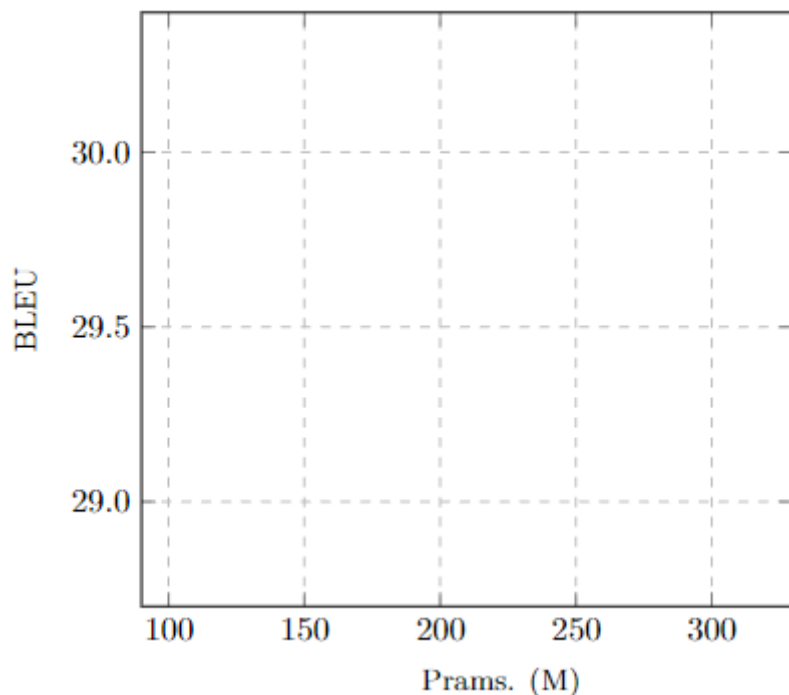
是axis命令的基本使用格式，与figure，tikz等命令类似；at={(0,0)}表示的是画布位于(0, 0)处；width = 0.7\textwidth, height=0.5\textwidth表示该画布的宽和高分别为0.7\textwidth 和 0.5\textwidth；xlabel与ylabel则为x轴与y轴的标签；坐标轴取值范围以及坐标的设置则由如下代码定义：

```
ymin=28.7,ymax=30.4, ytick={29,29.5,30}, xmin=90,xmax=330,xtick={100, 150, 200, 250, 300}
```

该命令将x轴的取值范围设置为[90, 330]；y轴的取值范围设置为[28.7, 30.4]；y轴的刻度线为{29,29.5,30}；x轴的刻度线为{100, 150, 200, 250, 300}。此外axis还提供了对坐标轴数值精度的设置，如下所示：

```
yticklabel style={/pgf/number format/precision=1, /pgf/number format/fixed zerofill}
```

该命令将精度设置为小数点后一位。



Step 2: 将坐标点置于画布之上。这一部分我们主要采用的是addplot命令，具体如下：

```
\begin{axis}[...]
  \addplot[purple!30,mark=*,mark size=3pt,thick,mark options=
{fill=purple!30,draw=purple!30,line width=1.0pt}] coordinates { (118,29.05)};

  \addplot[kleinblue!50,mark=*,mark size=6pt,thick,mark options=
{fill=kleinblue!70,draw=kleinblue!70,line width=1.0pt}] coordinates {
(194,29.44)};

  \addplot[gray!30,mark=*,mark size=4.5pt,thick,mark options=
{fill=gray!30,line width=1.0pt}] coordinates { (137,29.30)};

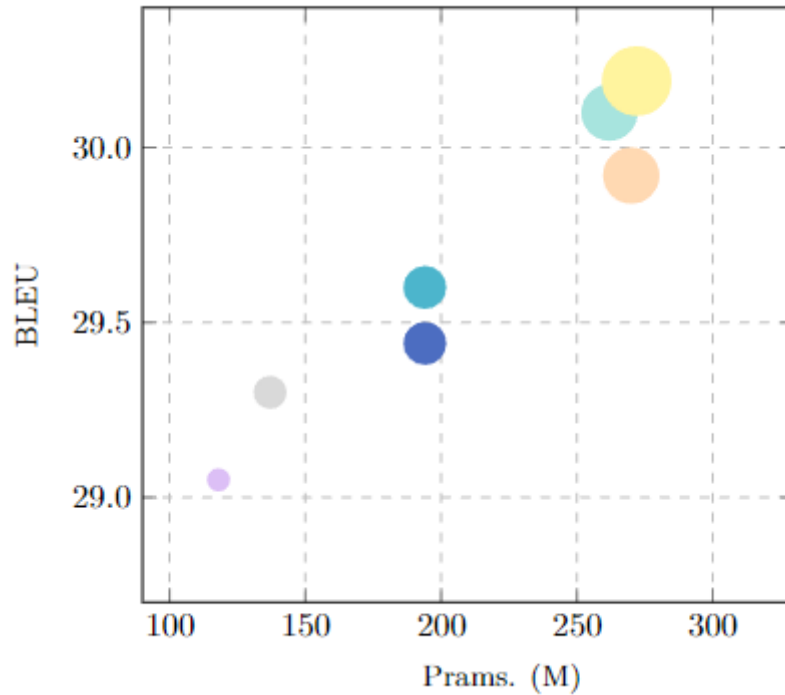
  \addplot[bangdibblue!70,mark=*,mark size=6pt,thick,mark options=
{fill=bangdibblue!70,line width=1.0pt}] coordinates { (194,29.60)};

  \addplot[orange!30,mark=*,mark size=8pt,thick,mark options=
{fill=orange!30, draw=orange!30,line width=1.0pt}] coordinates { (270,29.92)};

  \addplot[tiffanyblue!70,mark=*,mark size=8pt,thick,mark options=
{fill=tiffanyblue!70, draw=tiffanyblue!70,line width=1.0pt}] coordinates {
(262,30.10)};

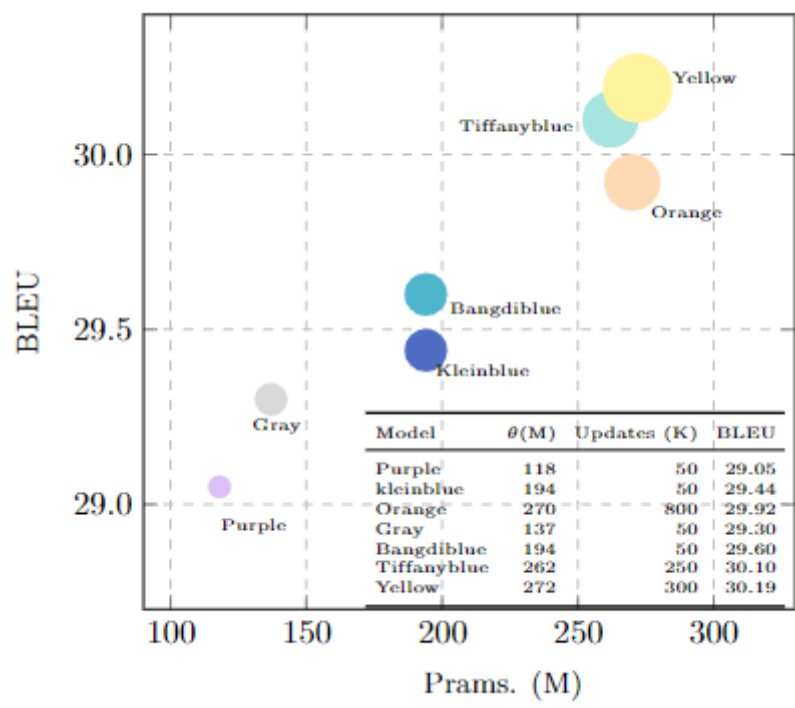
  \addplot[yellow!50,mark=*,mark size=10pt,thick,mark options=
{fill=yellow!50, draw=yellow!50,line width=1.0pt}] coordinates { (272,30.19)};
\end{axis}
```

\addplot命令的核心部分是 coordinates {}，正常情况下，我们只需要将对应的点坐标按照 (x y) 的格式进行填写即可。其中 \addplot[params] 的 params主要用于调整点的尺寸、点填充颜色以及点的形状。



Step 3: 绘制annotations以及表格的嵌入。这一部分主要采用的核心代码是基于node命令的。具体来说，我们只需要将node命令中{}中填充annotations名称或者表格即可。这里不再赘述，只提供相应的代码，如下：

```
\node[font=\tiny] at (3.3em, 2.5em){Purple};
\node[font=\tiny] at (4em, 5.5em){Gray};
\node[font=\tiny] at (10.9em, 9em){Bangdiblue};
\node[font=\tiny] at (10.2em, 7.2em){Kleinblue};
\node[font=\tiny] at (11.2em, 14.5em){Tiffanyblue};
\node[font=\tiny] at (16.9em, 16em){Yellow};
\node[font=\tiny] at (16.35em, 11.9em){Orange};
\node[] at (12.85em, 3.5em){
\setlength{\tabcolsep}{2.7pt}
\tiny
\begin{tabular}{lrrrr}
\toprule
Model & $\mathrm{\theta}$ (M) & Updates (K) & BLEU \\
\midrule
Purple & 118 & 50 & 29.05 \\
Kleinblue & 194 & 50 & 29.44 \\
Orange & 270 & 800 & 29.92 \\
Gray & 137 & 50 & 29.30 \\
Bangdiblue & 194 & 50 & 29.60 \\
Tiffanyblue & 262 & 250 & 30.10 \\
Yellow & 272 & 300 & 30.19 \\
\bottomrule
\end{tabular}
};
```



总结

Latex 绘图的核心还是node的使用，只要熟练掌握node的使用，任何复杂的图都可以轻松解决。当然，为了更加高效、简洁的绘制图形，我们经常会用到偏移方法来进行绘制，以便更好的建模不同图形之间的关系，为后续图形的改动提供便利条件。