Data split best practice:

<table>
<thead>
<tr>
<th>Labeled data</th>
<th>Train</th>
<th>Val</th>
<th>Test</th>
</tr>
</thead>
</table>

What %? Depends on size of data and amount of noise.
Smaller $\rightarrow$ higher variance
Larger $\rightarrow$ less data for other splits

For small data, take full advantage of as much data as possible:

| Val, Val, Val, ... Val, Test |

"k-fold cross-validation":

- train on each subset of k-1 chunks, val on the last
- avg val accuracy across all k trials
  + better training, lower-variance val accuracy
  - need to train k times

"leave-one-out cross-validation":

$k = n$
\[ S = \text{fit} \text{ transform} \]

\[
\begin{array}{c|c|c}
\text{Train} & \text{Val} & \text{Test} \\
\end{array}
\]

\[ S = \text{fit} \text{ transform } ([\text{train}]) \]

\[ S = \text{transform } ([\text{Val}]) \]

\[ S = \text{transform } ([\text{test}]) \]