The exam will be closed-book, closed-notes, and you will not be allowed to have anything other than the exam and a pen/pencil and an eraser on your desk. You will not get any scratch paper, so please get used to using the margins and backs of the exam paper itself as you practice. The amount of space provided under each question is not an indication of the length of the answer. Materials covered are all the materials in sections 0-2 of the class syllabus page and the first four homework assignments. Materials on the syllabus page include class notes and all the supplementary materials except those listed in parentheses. Appendices of the class notes are not required reading, except to the extent that they help you understand materials in the main text.

1. The 'valid' version of correlation for an image $I$ and template $T$ returns an output image that is defined at all and only the template positions for which $I$ and $T$ overlap completely.

Give the 'valid' normalized cross-correlation image $C$ of the $1 \times 3$ image

$$I = \begin{bmatrix} 3\sqrt{2} & \sqrt{2} & 0 \end{bmatrix}$$

and the $1 \times 2$ template

$$T = \begin{bmatrix} \sqrt{2} & -\sqrt{2} \end{bmatrix}.$$

Values above are chosen to make the math simple, so you do not need a calculator. If you are unsure about your manipulations, you may want to hand in your calculations (arranged neatly and legibly) for partial credit.
2. Give an expression for the derivative $f'(x)$ of the Gaussian function

$$f(x) = e^{-x^2/2}$$

and draw a qualitatively correct sketch a plot of $f'(x)$ on the Cartesian diagram below.

3. Write the numerical values of the orientation $\theta(x, y)$ of the gradient of the following function:

$$f(x, y) = e^{-\frac{1}{2}(x^2+y^2)}$$

for the values of $x$ and $y$ in the table below. (Just fill in the last column of the table). The values of $\theta(x, y)$ are angles in $[-\pi, \pi)$ or, if you prefer, $[-180, 180)$.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>$\theta(x, y)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>
4. The figure on the left below shows the samples of a discrete signal \( s(n) \) that is zero for all \( n < 0 \) and all \( n > 3 \). For \( n = 0, 1, 2, 3 \), the signal takes on values 1, 1, 3, 1. The figure on the right shows the graph of a continuous function \( p(x) \) defined as follows:

\[
p(x) = \begin{cases} 
  1/2 & \text{for } -1 \leq x \leq 0 \\
  1 & \text{for } 0 \leq x \leq 1 \\
  0 & \text{elsewhere.}
\end{cases}
\]

On the blank diagram below, draw an accurate plot of the hybrid convolution

\[
h(x) = \sum_{n=-\infty}^{\infty} s(n)p(x-n)
\]

of \( s \) and \( p \). Make sure you draw your plot for all values of \( x \) between -3 and 8 as a well-visible line. To keep your answer clean, you may want to use the figure above, left (diagram for \( s(n) \)) for scratch work, and then copy the result to the blank diagram.
5. The following image $J$ is the integral image of some other image $I$. What is $I$? Remember that $J$ has one more row and column than $I$.

$$J = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 2 & 3 \\ 0 & 3 & 6 & 11 \end{bmatrix}$$

6. Let $(x, y)$ be a point with real coordinates such that

$$a \leq x < a + 1 \quad \text{and} \quad b \leq y < b + 1$$

for integers $a$ and $b$. Also let

$$\Delta x = x - a \quad \text{and} \quad \Delta y = y - b.$$ 

Bilinear interpolation yields an image value at $(x, y)$ that has the following format:

$$I(x, y) = I(a, b) \quad \text{and} \quad \Delta x = x - a \quad \text{and} \quad \Delta y = y - b.$$ 

Fill in the missing parts in this formula.

7. Computation of a Gaussian pyramid applies an operator we called $\text{down}$ repeatedly to obtain each level of the pyramid from the previous level. What two operations does $\text{down}$ perform, and in what order?
8. What convolution kernel do Dalal and Triggs use to compute image derivatives in the horizontal direction? [Warning: convolution, not correlation.]

9. What is the key difference between regression and classification?

10. Define the misclassification loss function, either in words or by a formula.

11. Summarize the comparison of linear learners and decision trees in terms of bias that Pedro Domingo made in his article *A few useful things to know about machine learning*. Specifically, briefly state which has lower bias and why.
12. In his article *A few useful things to know about machine learning*, Pedro Domingo shows a bare-bones decision tree learner. What impurity measure does he use? Just give the name, you need not explain.

13. Add the most appropriate missing word to this sentence: A learner that overfits is said not to ________ well.

14. Give your best estimate of the Recall-Precision Equal-Error Rate (RPEER) for the diagram below.

15. If you want to increase the precision of a detector, would you increase or decrease the number of false positives?
16. You are training a classification tree for a binary labeling problem, using the empirical error as a measure of impurity. Features are real numbers (one-dimensional features), and possible labels are 0 and 1. The training samples that reach node \( n \) in the tree are as follows:

\[ S = \{(0, 1), (1, 0), (2, 1), (3, 1), (4, 1), (5, 0), (6, 0), (7, 0), (8, 1), (9, 1)\} \]

These samples are also drawn in the diagram below for your convenience.

(a) What is the impurity \( i(S) \) of \( S \)?

(b) Give the optimal splitting threshold \( \tau \) for \( S \).

(c) Training samples that are below the splitting threshold are placed in set \( L \) and the others in set \( R \). What is the combined impurity

\[ i_c = \frac{|L|}{|S|} i(L) + \frac{|R|}{|S|} i(R) \]

after the split? Show your calculations.

(d) Is it possible to improve purity in \( L \) and \( R \) by further splits? If you are sure, just answer yes or no.
17. Give an expression for the exact numerical value of the expected fraction of training data that are not used to train tree number 1 in a random forest, assuming that the training set \( T \) is infinitely large (that is, give the requested value as \( |T| \to \infty \)).

18. When is a training patch labeled as positive in a Hough forest?

19. Give the three matrices \( U, \Sigma, V \) of the full singular value decomposition \( A = U\Sigma V^T \) of the following matrix

\[
A = \begin{bmatrix}
1 & 0 \\
0 & -3 \\
0 & 0
\end{bmatrix}.
\]

The answer is unique up to some of the signs. [Hint: This problem can be solved by inspection, with essentially no calculations. Remember that to permute the rows of a matrix you pre-multiply it by a permutation of the identity matrix (i.e., a permutation matrix). Similarly for permuting columns, but then you post-multiply instead. Permutation matrices are trivially orthogonal.]
20. The following equalities hold for a certain matrix $A$:

\[ a = \arg \max_{\|x\|=1} Ax = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad b = \arg \min_{\|x\|=1} Ax = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \]

\[ Aa = \sqrt{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad Ab = 3\sqrt{2} \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}. \]

Give the three matrices $U$, $\Sigma$, $V$ of the full singular value decomposition $A = U \Sigma V^T$ of $A$. The answer is unique up to some sign combinations.